TRANSFORMED PLANT WITH MODIFIED ENVIRONMENTAL STRESS TOLERANCE GENE EXPRESSION

Inventors: Michael F. Thomashow, Eric J. Stockinger and Cai-Zhong Jiang

RELATIONSHIP TO COPENDING APPLICATION

This application is a continuation-in-part of U.S. Application Serial No.: 08/706,270; filed: September 4, 1996, entitled "DNA AND ENCODED PROTEIN WHICH REGULATES COLD AND DEHYDRATION REGULATED GENES" which is incorporated by reference.

FIELD OF THE INVENTION

The present invention relates to the regulatory response of plants to environmental stresses such as cold and to drought. More specifically, the present invention relates to genes which regulate the response of a plant to environmental stresses such as cold or drought and their use to enhance the stress tolerance of recombinant plants into which these genes are introduced.

BACKGROUND OF THE INVENTION

Environmental factors serve as cues to trigger a number of specific changes in plant growth and development. One such factor is low temperature. Prominent examples of cold-regulated processes include cold acclimation, the increase in freezing tolerance that occurs in response to low non-freezing temperatures (Guy, C. L., Annu. Rev. Plant Physiol. Plant Mol. Biol. **41**:187-223 (1990)); vernalization, the shortening of time to flowering induced by low temperature (Lang, A., in Encyclopedia of Plant Physiology, Vol. 15-1, ed. Ruhland, W. (Springer, Berlin), pp. 1489-1536 (1965)); and stratification, the breaking of seed dormancy by low temperature (Berry, J. A. and J. K. Raison, in Encyclopedia of Plant Physiology, Vol. 12A, eds. Lange, O. L., Nobel, P. S., Osmond, C. B. and Ziegler, H. (Springer, Berlin), pp. 277-338 (1981)). Due to the fundamental nature and agronomic importance of these processes, there is interest in understanding how plants sense and respond to low temperature. One approach being taken is to determine the signal transduction pathways and regulatory mechanisms involved in cold-regulated gene expression.

Strong evidence exists for calcium having a role in low temperature signal transduction and regulation of at least some COR (cold-regulated) genes. Dhindsa and colleagues (Monroy, A. F., et al, Plant Physiol. **102**:1227-1235 (1993); Monroy, A. F., and R. S., The Plant Cell, **7**:321-331 (1995)) have shown that, in alfalfa, calcium chelators and calcium channel blockers prevent low temperature induction of COR genes and that calcium ionophores and calcium channel agonists induce expression of COR genes at normal growth temperatures. Similarly, Knight et al (The Plant Cell **8**:489-503 (1996)) have shown that cold-induced expression of the *Arabidopsis thaliana* COR gene *KIN1* is inhibited by calcium chelators and calcium channel blockers. These results suggest that low temperature triggers an influx of extracellular calcium that activates a signal transduction pathway that induces the expression of COR genes. Consistent with this notion is the finding that low temperature evokes transient increases in cytosolic calcium levels in plants (Knight, M. R. et al, Nature **352**:524-526 (1991); Knight, H., et al., The Plant Cell **8**:489-503 (1996)). In addition, low temperatures have been shown to stimulate the activity of mechanosensitive calcium-selective cation channels in plants (Ding, J. P. and B. G. Pickard, Plant J. **3**:713-720 (1993)).

Recent efforts have led to the identification of a *cis*-acting cold-regulatory element in plants, the C-repeat/DRE (Yamaguchi-Shinozaki, et al., The Plant Cell **6**:251-264 (1994); Baker, S. S., et al., Plant. Mol. Biol. **24**:701-713 (1994); Jiang, C., et al., Plant Mol. Biol. **30**:679-684 (1996)). The element, which has a 5 base pair core sequence for CCGAC, is present once to multiple times in all plant cold-regulated promoters that have been described to date; these include the promoters of the *COR15a* (Baker, S. S., et al, Plant. Mol. Biol. **24**:701-713 (1994)), *COR78/RD29A* (Horvath, D. P., et al, Plant Physiol. **103**:1047-1053 (1993); Yamaguchi-Shinozaki, K., et al., The Plant Cell **6**:251-264 (1994)), *COR6.6* (Wang, H., et al., Plant Mol. Biol. **28**:605-617 (1995)) and *KIN1* (Wang, H., et al, Plant Mol. Biol. **28**:605-617 (1995)) genes of Arabidopsis and the *BN115* gene of *Brassica napus* (White, T. C., et al, Plant Physiol. **106**:917-928 (1994)). Deletion analysis of the Arabidopsis *COR15a* gene suggested that the CCGAC sequence, designated the C-repeat, might be part of a *cis*-acting cold-regulatory element (Baker, S. S., et al., Plant Mol. Biol. **24**:701-713 (1994)). That this was the case was first

demonstrated by Yamaguchi-Shinozaki and Shinozaki (Yamaguchi-Shinozaki, K., et al., The Plant Cell 6:251-264 (1994)) who showed that two of the C-repeat sequences present in the promoter of *COR78/RD29A* induced cold-regulated gene expression when fused to a reporter gene. It was also found that these two elements stimulate transcription in response to dehydration and high salinity and thus, was designated the DRE (dehydration, low temperature and high salt regulatory element). Recent studies by Jiang et al (Jiang, C., et al., Plant Mol. Biol. 30:679-684 (1996)) indicate that the C-repeats (referred to as low temperature response elements) present in the promoter of the *B. napus BN115* gene also impart cold-regulated gene expression.

U.S. Patents Nos. 5,296,462 and 5,356,816 to Thomashow describe the genes encoding the proteins involved in cold adaptation in *Arabidopsis thaliana*. In particular the DNA encoding the COR15 proteins is described. These proteins are significant in promoting cold tolerance in plants.

A need exists for the identification of genes which regulate the expression of cold tolerance genes and drought tolerance genes. A further need exists for DNA constructs useful for introducing these regulatory genes into a plant in order to cause the plant to begin expressing or enhance their expression of native or non-native cold tolerance genes and drought tolerance genes. These and other needs are provided by the present invention.

SUMMARY OF THE INVENTION

DNA in isolated form is provided which includes a sequence encoding a binding protein capable of selectively binding to a DNA regulatory sequence which regulates expression of one or more environmental stress tolerance genes in a plant. The binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in a plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes.

DNA in isolated form is also provided which includes a promoter and the sequence encoding the binding protein. In one variation, the promoter is an inducible promoter and/or is not induced by environmental stress.

A nucleic acid construct capable of transforming a plant is also provided which includes a sequence encoding a binding protein capable of selectively binding to a DNA regulatory sequence which regulates expression of one or more environmental stress tolerance genes in a plant. The binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in a plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes. The nucleic acid construct may be an RNA or DNA construct. Examples of types of constructs include, but are not limited to DNA and RNA viral vectors and plasmids.

A nucleic acid construct capable of transforming a plant is also provided which includes a sequence which when transformed into a plant expresses a binding protein capable of selectively binding to a DNA regulatory sequence which regulates one or more environmental stress tolerance genes in the plant. The binding protein preferably regulates expression of one or more environmental stress tolerance genes in the plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes.

In one variation of the above constructs, the construct also includes a promoter which regulates expression of the binding protein encoding sequence. The promoter may optionally be homologous or heterologous relative to the binding protein encoding sequence. The promoter and binding protein encoding sequence may also optionally be native to the same or a different plant species. In one variation, the promoter is an inducible promoter and/or is not induced by an environmental stress.

A binding protein in isolated form is also provided which is capable of selectively binding to a DNA regulatory sequence which regulates expression of one or more environmental stress tolerance genes in a plant. The binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in the plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes.

A recombinant binding protein expressed within a plant is also provided which is capable of selectively binding to a DNA regulatory sequence in the plant which regulates expression of one or more environmental stress tolerance genes in the plant. The recombinant binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in the plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes. The recombinant binding protein may be native or non-native to the plant. Further, the recombinant binding protein may be homologous or heterologous relative to the DNA binding protein present in the plant in which the binding protein is expressed.

A transformed cell of an organism is also provided which includes a recombinant sequence encoding a binding protein capable of selectively binding to a DNA regulatory sequence which regulates expression of one or more environmental stress tolerance genes in a plant. The binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in a plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes. The transformed cell may be a unicellular organism such as a bacterium, yeast or virus, or from a multicellular organism such as a fungus or a plant.

A transformed cell is also provided which includes a promoter and a sequence encoding a binding protein where at least one of the promoter and sequence under regulatory control of the promoter is recombinant. Optionally, one or both of the promoter and sequence under regulatory control of the promoter is not native to the cell. In one variation, the promoter is an inducible promoter and/or is not induced by environmental stress.

A transformed cell is also provided which includes a recombinant binding protein expressed within the cell which is capable of selectively binding to a DNA regulatory sequence in the plant which regulates expression of one or more environmental stress tolerance genes in the plant. The binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in the plant by selectively binding to a DNA regulatory sequence which

regulates the one or more environmental stress tolerance genes. The binding protein may be native or non-native to the cell.

A transformed plant with modified environmental stress tolerance gene expression is also provided. In one embodiment, the transformed plant includes one or more environmental stress tolerance genes; a DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes; and a recombinant sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence.

In another embodiment, the transformed plant includes one or more environmental stress tolerance genes; a DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes; a sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence; and a recombination promoter which regulates expression of the sequence encoding the binding protein.

In yet another embodiment, the transformed plant includes one or more environmental stress tolerance genes; a recombinant DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes; and a sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence.

In yet another embodiment, the transformed plant includes at least one recombinant environmental stress tolerance gene; a DNA regulatory sequence which regulates expression of the at least one environmental stress tolerance gene; and a sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence.

In yet another embodiment, the transformed plant includes at least one recombinant environmental stress tolerance gene; a DNA regulatory sequence which regulates expression of the at least one environmental stress tolerance gene; and a recombinant binding protein expressed by the plant which is capable of selectively binding to the DNA regulatory sequence.

In each of the above embodiments, the binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in a plant by selectively binding to a DNA regulatory sequence which regulates the one

or more environmental stress tolerance genes. The environmental stress tolerance gene, DNA regulatory sequence, and sequence encoding the binding protein may each independently be native or non-native to the plant and may each independently be homologous or heterologous relative to each other.

A method for altering an environmental stress tolerance of a plant is also provided. In one embodiment, the method includes transforming a plant with at least one copy of a gene encoding a binding protein capable of binding to a DNA regulatory sequence which regulates one or more environmental stress tolerance genes in the plant; expressing the binding protein encoded by the gene; and stimulating expression of at least one environmental stress tolerance gene through binding of the binding protein to the DNA regulatory sequence.

In another embodiment, the method includes transforming a plant with a promoter which regulates expression of at least one copy of a gene encoding a binding protein capable of binding to a DNA regulatory sequence which regulates one or more environmental stress tolerance genes in the plant; expressing the binding protein encoded by the gene; and stimulating expression of at least one environmental stress tolerance gene through binding of the binding protein to the DNA regulatory sequence.

In another embodiment, the method includes transforming a plant with one or more environmental stress tolerance genes whose expression is regulated by a DNA regulatory sequence; and expressing a binding protein capable of binding to the DNA regulatory sequence and activating expression of the one or more environmental stress tolerance genes.

According to any one of the above embodiments of the present invention, the binding protein may optionally be selected such that it selectively binds to a member of a class of DNA regulatory sequences which includes the subsequence CCG or more particularly one of the following subsequences: CCGAA, CCGAT, CCGAC, CCGAG, CCGTA, CCGTT, CCGTC, CCGTG, CCGCA, CCGCT, CCGCG, CCGCC, CCGGA, CCGGT, CCGGC, CCGGG, AACCG, ATCCG, ACCCG, AGCCG, TACCG, TTCCG, TCCCG, TGCCG, CACCG CCCG, GACCG, GTCCG, GCCCG, GGCCG, ACCGA, ACCGT, ACCGC, ACCGG, TCCGA, TCCGT, TCCGC, TCCGG, CCCGA,

CCCGT, CCCGC, CCCGG, GCCGA, GCCGT, GCCGC, and GCCGG. The binding protein may also be selected to include an AP2 domain.

Optionally, the sequence encoding the binding protein satisfies one or more of the following limitations: it is a homolog of the AP2 binding domain of CBF1, CBF2, or CBF3; it is a homolog of CBF1, CBF2, or CBF3; it is a homolog of CBF1, CBF2 or CBF3 which includes a sequence which has at least about 83 percent identity to at least one of CBF1, CBF2, and CBF3; it encodes an amino acid sequence which has substantially the same homology to CBF1, CBF2, and CBF3 as these amino acid sequences have with each other; it includes one of the DNA sequences listed in SEQ. ID. NOs. 1, 12 and 14; and it encodes a protein having one of the amino acid sequences listed in SEQ. ID. NOs. 2, 13 and 15. The sequence encoding the binding protein may be a naturally occuring sequence native to a plant or a non-naturally occuring sequence. For example, the sequence may include a subsequence which encodes a binding domain for the DNA regulatory sequence fused to a transcription activating region, such as the transcription activating region of VP16 or GAL4.

A method is also provided for identifying from a cDNA library of at least a portion of a plant genome a gene sequence encoding a protein capable of binding to a target DNA regulatory sequence. In one embodiment, the method comprises

taking a microorganism which includes a target DNA regulatory sequence for one or more environmental stress tolerance genes, a transciption activator for activating expression of a reporter gene, and a reporter gene whose expression is activated by a protein which includes a binding domain capable of binding to the target DNA regulatory sequence and an activation domain capable of activating the transcription activator;

fusing sequences from a cDNA library of at least a portion of a plant genome to a sequence which encodes a functional activation domain in the microorganism;

introducing the fused sequences into the microorganism; and

selecting microorganisms which express the reporter gene, expression of the reporter gene indicating expression of a fusion

protein which includes a binding domain for the target DNA regulatory sequence and theactivation domain; and

identifying the gene sequence from the cDNA library introduced into the microorganism.

The target DNA regulatory sequence may optionally include the subsequence CCG or the subsequence CCGAC. This embodiment of the invention also relates to DNA in substantially isolated form, nucleic acid constructs capable of transforming a plant, cells, and transformed plants which include a gene sequence identified by this method.

These and other objects will become increasingly apparent by reference to the following description and the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B show how the yeast reporter strains were constructed.

Figure 1A is a schematic diagram showing the screening strategy.

Figure 1B is a chart showing activity of the "positive" cDNA clones in yeast reporter strains.

Figures 2A, 2B, 2C and 2D provide an analysis of the pACT-11 cDNA clone.

Figure 2A is a schematic drawing of the pACT-11 cDNA insert indicating the location and 5' to 3' orientation of the 24 kDa polypeptide and 25s rRNA sequences.

Figure 2B is a DNA and amino acid sequence of the 24 kDa polypeptide (SEQ ID NO:1 and SEQ ID NO:2).

Figure 2C is a schematic drawing indicating the relative positions of the potential nuclear localization signal (NLS), the AP2 domain and the acidic region of the 24 kDa polypeptide.

Figure 2D is a chart showing comparison of the AP2 domain of the 24 kDa polypeptide with that of the tobacco DNA binding protein EREBP2.

Figure 3 is a chart showing activation of reporter genes by the 24 kDa polypeptide.

Figure 4 is a photograph of an electrophoresis gel showing expression of the recombinant 24 kDa polypeptide in *E. coli.*

Figure 5 is a photograph of a gel for shift assays indicating that CBF1 binds to the C-repeat/DRE.

Figure 6 is a photograph of a southern blot analysis indicating *CBF1* is a unique or low copy number gene.

Figures 7A, 7B and 7C relate to *CBF1* transcripts in control and cold-treated Arabidopsis.

Figure 7A is a photograph of a membrane RNA isolated from Arabidopsis plants that were grown at 22°C or grown at 22°C and transferred to 2.5°C for the indicated times.

Figure 7B is a graph showing relative transcript levels of *CBF1* in control and cold-treated plants.

Figure 7C is a graph showing relative transcript levels of *COR15a* in control and cold-treated plants.

Figure 8 is a Northern blot showing *CBF1* and *COR* transcript levels in RLD and transgenic *Arabidopsis* plants.

Figure 9 is an immunoblot showing *COR15a*m protein levels in RLD and transgenic *Arabidopsis* plants.

Figures 10A and 10B are graphs showing freezing tolerance of leaves from RLD and transgenic *Arabidopsis* plants.

Figure 11 is a photograph showing freezing survival of RLD and A6 *Arabidopsis* plants.

Figure 12 shows the DNA sequence for *CBF2* encoding CBF2.

Figure 13 shows the DNA sequence for *CBF3* encoding CBF3.

Figure 14 shows the amino acid alignment of proteins CBF1, CBF2 and CBF3.

Figure 15 is a graph showing transcription regulation of COR genes by *CBF1*, *CBF2* and *CBF3* genes in yeast.

Figure 16 shows the amino acid sequence of a canola homolog and its alignment to the amino acid sequence of CBF1.

DETAILED DESCRIPTION

The present invention relates to DNA encoding binding proteins capable of binding to a DNA regulatory sequence which regulates expression of one or more environmental stress tolerance genes in a plant. The present invention also relates to the binding proteins encoded by the DNA. The DNA and binding proteins may be native or non-native relative to the DNA regulatory sequence of the plant. The DNA and binding proteins may also be native or non-native relative to environmental stress tolerance genes of the plant which are regulated by the DNA regulatory sequence.

The present invention also relates to methods for using the DNA and binding proteins to regulate expression of one or more native or non-native environmental stress tolerance genes in a plant. These methods may include introducing DNA encoding a binding protein capable of binding to a DNA regulatory sequence into a plant, introducing a promoter into a plant which regulates expression of the binding protein, introducing a DNA regulatory sequence into a plant to which a binding protein can bind, and/or introducing one or more environmental stress tolerance genes into a plant whose expression is regulated by a DNA regulatory sequence.

The present invention also relates to recombinant plants and plant materials (e.g., plant tissue, seeds) into which one or more gene sequences encoding a binding protein have been introduced as well as plants and plant materials within which recombinant binding proteins encoded by these gene sequences are expressed. By introducing a gene sequence encoding a binding protein into a plant, a binding protein can be expressed within the plant which regulates expression of one or more stress tolerance genes in the plant. Regulation of expression can include causing one or more stress tolerance genes to be expressed under different conditions than it would be in the plant's native state, increasing a level of expression of one or more stress tolerance genes, and/or causing the expression of one or more stress tolerance genes to be inducible by an exogenous agent. Expression of the binding protein can be under the control of a variety of promoters. For example, promoters can be used to overexpress the binding protein, change the environment conditions under which the binding protein is expressed, or enable the expression of the binding protein to be induced, for example by the addition of an exogenous inducing agent.

The present invention also relates to recombinant plants and plant materials into which a recombinant promoter is introduced which controls a level of expression of one or more gene sequences encoding a binding protein. The one or more gene sequences may be recombinant native or non-native sequences or may be native, non-recombinant gene sequences whose expression is altered by the introduction of the recombinant promoter.

The present invention also relates to recombinant plants and plant materials into which a recombinant native or non-native DNA regulatory sequence is introduced which regulates expression of one or more native or non-native environmental stress tolerance genes.

Examples of environmental stresses for which stress tolerance genes are known to exist include, but are not limited to, cold tolerance, dehydration tolerance, and salinity tolerance. As used herein, environmental stress tolerance genes refer to genes which function to acclimate a plant to an environment stress. For example, cold tolerance genes, also referred to as COR genes (COld Regulated), refer to genes which function to acclimate a plant to a cold temperature environment. These genes typically are activated when a plant is exposed to cold temperatures. Dehydration tolerance genes refer to genes which function to acclimate a plant to dehydration stress. These genes typically are activated in response to dehydration conditions which can be associated with drought or cold temperatures which cause water in the plant to freeze and thereby dehydrate the plant tissue. It is noted that some cold tolerance genes may function to provide a plant with a degree of dehydration tolerance and visa versa. For example, COR genes are known to also be activated by dehydration stress. This application is intended to encompass genes which regulate one or more environmental stress tolerance genes such as cold tolerance genes, dehydration tolerance genes, and genes which perform a dual function of cold and dehydration tolerance.

One embodiment of the invention relates to a DNA sequence in isolated form which includes a sequence which can selectively bind to a DNA regulatory sequence which regulates one or more environmental stress tolerance genes. The binding protein encoded by the DNA sequence preferably also can regulate expression of one or more environmental stress tolerance genes in a plant by its selective binding

to a DNA regulatory sequence. The DNA sequence may exist in a variety of forms including a plasmid or vector and can include sequences unrelated to the gene sequence encoding the binding protein. For example, the DNA sequence can include a promoter which regulates expression of the regulatory gene.

In one variation of this embodiment, the DNA regulatory sequence is a C-repeat cold and drought regulation element (C-repeat/DRE). As will be explained and demonstrated herein, C-repeat/DRE regulatory sequences appear to be conserved in plants with some degree of variability plant to plant. Using the teachings of the present invention, C-repeat/DRE regulatory sequences native to different plants can be identified as well as the native stress tolerance regulatory genes which encode for proteins which bind to the C-repeat/DRE DNA regulatory sequences. Hence, although the examples provided herein to describe the present invention are described with regard to the *Arabadopsis* C-repeat/DRE DNA regulatory sequence, the present invention is not intended to be limited to the *Arabadopsis* C-repeat/DRE DNA regulatory sequence. Rather, the *Arabadopsis* C-repeat/DRE DNA regulatory sequence is believed to be a member of a class of binding domains which includes the subsequence CCGAC which in turn is believed to be a member of a class of binding domains which includes the subsequence CCG. Other, different classes of binding domains may also exist. The teachings of the present invention may be used to identify sequences which bind to these and other classes of binding domains once they are identified.

In one variation of this embodiment, the gene sequence encodes a binding protein which selectively binds to a member of a class of DNA regulatory sequences which includes the subsequence CCG. In another variation, the gene sequence encodes a binding protein which selectively binds to a member of a class of DNA regulatory sequences which includes the subsequence CCGAC. The CCGAC subsequence has been found to present in the C-repeat/DRE DNA regulatory sequences of *Arabadopsis* and *Brassica* and to function in Tobacco based on the ability of the C-repeat/DRE to direct cold and tolerance regulated gene expression.

In yet another variation, the stress tolerance regulatory gene sequence encodes a binding protein which includes an AP2 domain. It is believed that a significant class of environmental stress tolerance regulatory genes encode for

binding proteins with an AP2 domain capable of binding to the DNA regulatory sequence. The AP2 domain of the binding protein is preferably a homolog of the AP2 domain of one of the CBF binding proteins described herein. The subsequence encoding the AP2 domain is preferably a homolog of a subsequence of one of the CBF genes described herein which encodes an AP2 domain.

In another variation of this embodiment, the gene sequence is a homolog of CBF1, CBF2, or CBF3. In one preferred embodiment, the gene sequence is a homolog of CBF1, CBF2, or CBF3 and has at least about 83 percent identity to CBF1, CBF2, or CBF3. In another preferred embodiment, the gene sequence encodes an amino acid sequence which has substantially the same homology to CBF1, CBF2, and CBF3 as these amino acid sequences have with each other.

In yet another variation of this embodiment, the isolated DNA sequence includes one of the DNA sequences listed in SEQ. ID. NOs. 1, 12 and 14.

In yet another variation of this embodiment, the isolated DNA sequence encodes a protein having one of the amino acid sequences listed in SEQ. ID. NOs. 2, 13 or 15.

The present invention also relates to a method for identifying gene sequences from at least a portion of a plant genome which encode binding proteins capable of binding to a target DNA regulatory sequence which regulates expression of one or more stress tolerance genes in a plant.

In one embodiment, the method includes:

taking a microorganism which includes a target DNA regulatory sequence for one or more environmental stress tolerance genes, a transciption activator for activating expression of a reporter gene, and a reporter gene whose expression is activated by a protein which includes a binding domain capable of binding to the target DNA regulatory sequence and an activation domain capable of activating the transcription activator;

fusing sequences from a cDNA library of at least a portion of a plant genome to a sequence which encodes a functional activation domain in the microorganism;

introducing the fused sequences into the microorganism; and

selecting microorganisms which express the reporter gene, expression of the reporter gene indicating expression of a fusion protein which includes a binding domain for the target DNA regulatory sequence and theactivation domain; and

identifying the gene sequence from the cDNA library introduced into the microorganism.

In one variation of the method, the target DNA regulatory sequence includes the subsequence CCG and in another embodiment includes the subsequence CCGAC. In yet another variation, the target DNA regulatory sequence is the C-repeat/DRE for *Arabadopsis*. According to the above method, the target DNA regulatory sequence is preferably native to the plant family and more preferably to the plant species from which the cDNA library is derived.

In another variation of this embodiment, the cDNA library used in the method consists of sequences which encode for a protein having an AP2 domain since it is believed that a significant class of genes encoding binding proteins for stress tolerance genes encode an AP2 domain. As will be explained herein, screening for DNA sequences from a plant genome which exhibit this functional feature has been shown to be effective for isolating gene sequences encoding binding proteins of the present invention.

In another variation of this method, the sequences from the cDNA library are fused to a sequence which includes a selectable marker, the method further including the step of selecting for microorganisms expressing the selectable marker.

While the above methodology of the present invention is described herein with regard to identifying binding protein gene sequences from *Arabidopsis* cDNA using the C-repeat/DRE regulatory sequence for *Arabidopsis*, it is noted that this methodology can be readily used to identify regulatory binding protein gene sequences for other plants by using a DNA regulatory sequence native to those plants. Alternatively, different permutations of the CCG subsequence can be used as the target DNA regulatory sequence.

An example of a microorganism which may be used in the above method is yeast. cDNA can be introduced into the microorganism by a variety of mechanisms including plasmids and vectors. In one particular embodiment, the reporter gene is β-galactosidase.

The present invention also relates to any DNA sequences and binding proteins encoded by those DNA sequences which are identified by the above screening method.

The present invention also relates to a protein expressed by an environmental stress tolerance regulatory gene according to the present invention which can function *in vivo* in a plant to regulate expression of one or more environmental stress tolerance genes.

According to one embodiment, the protein is a recombinant binding protein expressed by a copy of a recombinant gene which is either not native to the plant or is native to the plant but introduced into the plant by recombinant methodology. For example, one might wish to introduce one or more copies of a regulatory gene which is native to the plant but is under the control of a promoter which overexpresses the binding protein, expresses the binding protein independent of an environmental stress, or can be induced to express the binding protein by an agent to which the plant can be exposed. Alternatively, one might wish to introduce one or more copies of a regulatory gene which is not native to the plant. For example, the non-native regulatory gene may be used to alter the way in which native environmental stress tolerance genes are regulated. Alternatively, the non-native regulatory gene may be used to regulate environmental stress tolerance genes which are also not native to the plant.

In another embodiment, the proteins have been isolated from a recombinant organism. The organism may be a microorganism (e.g., bacteria, yeast) or a multicellular organism such as a plant. In one variation, the protein is in substantially isolated form.

In yet another embodiment, the protein is a native, non-recombinant binding protein whose expression is regulated within a plant by a recombinant native or non-native promoter. For example, one might wish to replace a native promoter with a recombinant promoter which overexpresses the binding protein, expresses the binding protein independent of an environmental stress, or can be induced to express the binding protein by an agent to which the plant can be exposed.

In one variation of the above embodiments, the protein is capable of selectively binding to a DNA regulatory sequence for one or more environmental

stress tolerance genes in a plant. In another variation, the protein includes an AP2 domain which is capable of selectively binding to a DNA regulatory sequence for one or more environmental stress tolerance genes in a plant. One method which may be used to determine whether the protein binds selectively to the DNA regulatory sequence is a gel shift assay. The DNA regulatory sequence may optionally include a CCG subsequence, a CCGAC subsequence and optionally the C-repeat / DRE sequence of *Arabadopsis*.

In a particular embodiment, the binding protein is encoded by a DNA sequence which is a homolog of at least one of CBF1, CBF2, and CBF3. In one preferred embodiment, the DNA sequence encoding the binding protein has at least about 83 percent identity to at least one of CBF1, CBF2, and CBF3. In another preferred embodiment, the binding protein has an amino acid sequence with at least about 84 percent identity to at least one of CBF1, CBF2, and CBF3. In yet another preferred embodiment, the binding protein has an amino acid sequence which has substantially the same homology with CBF1, CBF2, or CBF3 as these amino acid sequences have with each other, as illustrated in Figure 14. In another embodiment, the binding protein has one of the amino acid sequences listed in SEQ. ID. NOs. 2, 13 or 15. In another embodiment, the binding protein is an isolated protein or a recombinantly produced protein which has a molecular weight of about 26 kDa as measured in an electrophoresis gel and binds to a DNA regulatory sequence which regulates a cold or dehydration regulated gene of *Arabidopsis thaliana*.

The present invention also relates to DNA and RNA constructs, such as plasmids, vectors, and the like, which are capable of transforming a plant. The constructs include a sequence which encodes a binding protein capable of selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes. The binding protein is preferably able to regulate expression of one or more environmental stress tolerance genes in a plant by selectively binding to the DNA regulatory sequence. More preferably, when transformed into a plant, the sequence regulates expression of one or more environmental stress tolerance genes in the plant by expressing the binding protein.

In one embodiment, the DNA construct includes a promoter and a regulatory gene sequence whose expression is under the control of the promoter. Different

promoters may be used to select the degree of expression or conditions under which the regulatory gene is expressed. For example, the promoter can be used to cause expression of the regulatory gene without an environmental stress or at a different level of environmental stress. The promoter may also be inducible, i.e., express the regulatory gene in response to an agent to which the promoter is exposed.

In another embodiment, the DNA construct includes one of the DNA sequences listed in SEQ. ID. NOs. 1, 12 and 14. In another embodiment, the DNA construct includes a DNA sequence which encodes for a protein having one of the amino acid sequences listed in SEQ. ID. NOs. 2, 13 or 15. The present invention also relates to plasmids pCBF1 (ATCC 98063), pCBF2 (ATCC ___), and pCBF3 (ATCC ___).

The present invention also relates to a recombinant microorganism, such as a bacterium, yeast, fungus, virus, into which at least one copy of a regulatory gene encoding a binding protein of the present invention has been introduced by a recombinant methodology.

The present invention also relates to recombinant plants into which at least copy of a regulatory gene encoding a binding protein of the present invention has been introduced by a recombinant methodology. The recombinant copy of the regulatory gene may be native or non-native to the plant and express a binding protein which is either native or non-native to the plant.

Expression of the recombinant copy of the regulatory gene may be under the control of the promoter. The promoter may increase the level at which the regulatory gene is expressed, express the regulatory gene without being induced by an environmental stress or express the regulatory gene in response to a different form or degree of environmental stress that would otherwise be needed to induce expression of the regulatory gene. For example, a promoter can be used which turns on at a temperature that is warmer than the temperature at which the plant normally exhibits cold tolerance. This would enable the cold tolerance thermostat of a plant to be altered. Similarly, a promoter can be used which turns on at a dehydration condition that is wetter than the dehydration condition at which the plant normally exhibits dehydration tolerance. This would enable the level at which a plant responds to dehydration to be altered. The promoter may also be inducible, i.e.,

express the regulatory gene in response to an agent to which the promoter is exposed. This would enable stress tolerance to be induced by applying an inducing agent to the plant.

The environmental stress tolerance gene regulated by the recombinantly expressed regulatory gene may be native or non-native to the plant. Hence, in one embodiment, the plant includes a recombinant copy of a regulatory gene which is native to the plant and expresses a native protein which functions within the plant to regulate expression of a native environmental stress tolerance gene. In this embodiment, the recombinant plant expresses a higher level of the native regulatory gene than the plant would otherwise.

In another embodiment, at least one of the regulatory genes and the environmental stress tolerance genes is not native to the plant. For example, the regulatory gene can be native and the environmental stress tolerance gene is non-native, or the regulatory gene is non-native and the environmental stress tolerance gene is native to the plant.

In yet another embodiment, the plant can include a recombinant copy of a regulatory gene which is not native to the plant as well as a recombinant copy of one or more environmental stress tolerance genes which also is not native to the plant. According to this embodiment, the non-native regulatory gene expresses a non-native binding protein which functions within the plant to regulate expression of the one or more non-native environmental stress tolerance genes. In this regard, it is envisioned that the present invention can be used to introduce, change and/or augment the environmental stress tolerance of a plant by introducing and causing the expression of environmental stress tolerance which the plant does not have in its native form. Accordingly, plants from warmer climates can be engineered to include one or more cold tolerance genes along with a regulatory gene needed to cause expression of the cold tolerance genes in the plant so that the engineered plant can survive better in a colder climate. Similarly, a plant can be engineered to include one or more dehydration tolerance genes along with a regulatory gene needed to cause expression of the dehydration tolerance gene so that the engineered plant can grow better in a dryer climate. In this regard, it should be possible to take a plant which grows well in a first climate and engineer it to include stress tolerance

genes and regulatory genes native to a second climate so that the plant can grow well in the second climate.

The present invention also relates to a method for changing or enhancing the environmental stress tolerance of a plant.

In one embodiment, the method includes introducing at least one copy of a regulatory gene encoding a binding protein of the present invention into a plant; expressing the binding protein encoded by the regulatory gene; and using the expressed binding protein to stimulate expression of at least one environmental stress tolerance gene through binding to a DNA regulatory sequence. According to this embodiment, the regulatory gene may be non-recombinant or recombinant native or non-native to the plant. Similarly, the DNA regulatory sequence and the environmental stress tolerance gene may each independently be native or non-native to the plant. In one variation of this embodiment, the method further includes recombinantly introducing an environmental stress tolerance gene into the plant which is regulated by the recombinant regulatory gene.

In another embodiment, the method includes introducing a recombinant promoter which regulates expression of a regulatory gene encoding a binding protein of the present invention into a plant; expressing the binding protein under the control of the recombinant promoter; and using the expressed binding protein to stimulate expression of at least one environmental stress tolerance gene through binding to a DNA regulatory sequence. According to this embodiment, the regulatory gene, the DNA regulatory sequence and the environmental stress tolerance gene may each independently be non-recombinant or recombinant native or non-native to the plant.

In yet another embodiment, the method includes introducing at least one recombinant environmental stress tolerance gene into a plant; expressing a binding protein; and using the expressed binding protein to stimulate expression of the recombinant environmental stress tolerance gene through binding to a DNA regulatory sequence. According to this embodiment, the gene encoding the regulatory protein, and the DNA regulatory sequence may each independently be non-recombinant or recombinant native or non-native to the plant. The recombinant environmental stress tolerance gene may be either native or non-native to the plant.

1. DEFINITIONS

The term "C-repeat cold and drought regulation element" or "C-repeat/DRE" refers to a sequence which includes CCG and functions as a binding domain in a plant to regulate expression of one or more environmental stress tolerance genes, such as cold or dehydration stress tolerance genes.

The term "cold stress" refers to a decrease in ambient temperature, including a decrease to freezing temperatures, which causes a plant to attempt to acclimate itself to the decreased ambient temperature.

The term "dehydration stress" refers to drought, high salinity and other conditions which cause a decrease in cellular water potential in a plant.

Transformation means the process for changing the genotype of a recipient organism by the stable introduction of DNA by whatever means.

A transgenic plant is a plant which contains DNA sequences which were introduced by transformation. Horticultural and crop plants particularly benefit from the present invention.

Translation means the process whereby the genetic information in an mRNA molecule directs the order of specific amino acids during protein synthesis.

The term "essentially homologous" means that the DNA or protein is sufficiently duplicative of that set forth in Figure 2B to produce the same result. Such DNA can be used as a probe to isolate DNA's in other plants.

A promoter is a DNA fragment which causes transcription of genetic material. For the purposes described herein, promoter is used to denote DNA fragments that permit transcription in plant cells.

A poly-A addition site is a nucleotide sequence which causes certain enzymes to cleave mRNA at a specific site and to add a sequence of adenylic acid residues to the 3'-end of the mRNA.

The phrase "DNA in isolated form" refers to DNA sequence which has been at least partially separated from other DNA present in its native state in an organism. A cDNA library of genomic DNA is not "DNA in isolated form" whereas DNA which

has been at least partially purified by gel electrophoresis corresponds to "DNA in isolated form".

2. C-Repeat/DRE Regulatory Elements In Plants

C-repeat cold and drought regulation elements (C-repeat/DRE) are sequences which function as a *cis*-acting regulatory element that stimulates transcription in response to an environmental stress, such as low temperature (Yamaguchi-Shinozaki, K., et al., The Plant Cell **6**:251-264 (1994); and Baker, S. S., et al., Plant Mol. Biol. **24**:701-713 (1994); Jiang, C., et al., Plant Mol. Biol. **30**:679-684 (1996)) or dehydration stress and high salinity (Yamaguchi-Shinozaki, K., et al., The Plant Cell **6**:251-264 (1994)). An object of the research leading to the present invention was the determination of how a C-repeat/DRE stimulates gene expression in response to these environmental factors, and whether cold, dehydration and high salinity affect independent or overlapping regulatory systems.

The first step toward determining how a C-repeat/DRE regulation element stimulates gene expression was the identification of the C-repeat cold and drought regulation element itself. The 5 base pair core sequence, CCGAC, has been found to be present once to multiple times in a variety of plant cold-regulated promoters in *Arabidopsis* and *Brassica* including the *COR15a* (Baker, S. S., et al, Plant. Mol. Biol. **24**:701-713 (1994)); *COR78/RD29A* (Horvath, D. P., et al, Plant Physiol. **103**:1047-1053 (1993) and Yamaguchi-Shinozaki, K., et al., The Plant Cell **6**:251-264 (1994)); *COR6.6* (Wang, H., et al., Plant Mol. biol. **28**:605-617 (1995)); and *KIN1* (Wang, H., et al, Plant Mol. Biol. **28**:605-617 (1995)) genes of Arabidopsis and the *BN115* gene of *Brassica napus* (White, T. C., et al, Plant Physiol. **106**:917-928 (1994)). As shown in the examples herein, core sequence CCGAC was used to identify proteins encoded by genes within the *Arabidopsis* genome which bind to this core sequence.

Applicants believe that the CCGAC core sequence is a member of family of core sequences having the common subsequence CCG. The binding of CBF1 to the C-repeat/DRE involves the AP2 domain. In this regard, it is germane to note that the tobacco ethylene response element, AGCCGCC, closely resembles the C-

repeat/DRE sequences present in the promoters of the Arabidopsis genes *COR15a*, GGCCGAC, and *COR78/RD29A*, TACCGAC. While the specific teachings in the present invention used only a DNA regulatory sequence which includes a CCGAC subsequence as the C-repeat/DRE core regulatory sequence, Applicants believe that other C-repeat/DRE regulatory sequences exist which belong to a broader CCG family of regulatory sequences. By screening plant genomes according to the methodology taught herein using other members of the CCG family, additional regulatory sequences as well as the binding proteins which bind to these regulatory sequences can be identified. For example, plants which are known to exhibit a form of environmental stress tolerance can be screened according to the blue colony assay and other screening methodologies used in the present invention with other members of the CCG family in order to identify other binding proteins and their gene sequences. Examples of other members of the CCG family include, but are not limited to, binding domains which include one of the following sequences: CCGAA, CCGAT, CCGAC, CCGAG, CCGTA, CCGTT, CCGTC, CCGTG, CCGCA, CCGCT, CCGCG, CCGCC, CCGGA, CCGGT, CCGGC, CCGGG, AACCG, ATCCG, ACCCG, AGCCG, TACCG, TTCCG, TCCCG, TGCCG, CACCG, CTCCG, CGCCG, CCCCG, GACCG, GTCCG, GCCCG, GGCCG,
ACCGA, ACCGT, ACCGC, ACCGG, TCCGA, TCCGT, TCCGC, TCCGG, CCCGA, CCCGT, CCCGC, CCCGG, GCCGA, GCCGT, GCCGC, and GCCGG.

Applicants also believe that other families of environmental stress tolerance DNA regulatory sequences, other than the CCG family may exist. The methodologies of the present invention may be used once such other families are identified in order to identify specific environmental stress tolerance DNA regulatory sequences and associated binding proteins.

3. Identification Of Environmental Stress Tolerance Regulatory Gene Sequences Using Target Regulatory Sequence

It is possible to take a cDNA library of at least a portion of a plant genome and screen the cDNA library for the presence of regulatory gene sequences which encode binding proteins capable of binding to a target regulatory sequence. As

used here, a target DNA regulatory sequence refers to a sequence to which a binding protein for one or more environmental stress tolerance genes binds. Permutions of the CCG and CCGAC families of DNA regulatory sequences represent examples of target DNA regulatory sequences. As detailed in Example 1 herein, this was the approach was used to identify CBF1, a sequence which encodes a binding protein for the *Arabadopsis* DNA regulatory sequence, from an *Arabadopsis* cDNA library .

First a target regulatory sequence is selected. The target regulatory sequence is preferably native to the plant from which the cDNA library being screened is derived.

Once a target regulatory sequence is selected, the target regulatory sequence is fused to a reporter gene and introduced into a microorganism. Expression of the reporter gene can be activated by a protein which includes a binding domain capable of binding to the target DNA regulatory sequence and an activation domain capable of activating transcription.

Sequences from a cDNA library of at least a portion of a plant genome are then fused to a sequence which encodes a functional activation domain in the microorganism. The fused sequences are then introduced into the microorganism. It is possible that the sequence from the cDNA library may already encode a functional activation domain, for example as described herein in Example 1.

Microorganisms which express the reporter gene are then selected. Since only those microorganisms which express a fusion protein which includes a binding domain for the target DNA regulatory sequence and an activation domain will stimulate expression of the reporter gene, expression of the reporter gene indicates expression of such a fusion protein.

The gene sequence from the cDNA library introduced into the microorganism which stimulates expression of the reporter gene is then identifed.

According to the above method, the target DNA regulatory sequence preferably includes the subsequence CCG and more preferably includes the subsequence CCGAC.

The "one-hybrid" strategy described in Li, J. J. and I. Herskowitz, Science **262**:1870-1874 (1993) and used in Example 1 to screen Arabidopsis cDNA is an

example of this method. This method can be used to screen any plant species for cDNAs that encode a target regulatory sequence, such as a C-repeat/DRE regulatory sequence. According to the "one hybrid" strategy, yeast strains are constructed that contain a *lacZ* reporter gene with either wild-type or mutant versions of target regulatory sequences in place of the normal UAS (upstream activator sequence) of the *GAL1* promoter. Yeast strains carrying these reporter constructs produce low levels of ß-galactosidase and form white colonies on filters containing X-gal. Reporter strains carrying wild-type target regulatory sequences are transformed with a cDNA expression library that contains random cDNA inserts fused to the acidic activator domain of the yeast GAL4 transcription factor "GAL4-ACT". Recombinant plasmids in the expression library that contain a cDNA insert encoding a C-repeat/DRE binding domain fused to GAL4-ACT will express fusion proteins which bind upstream of the *lacZ* reporter genes carrying the wild-type target regulatory sequence, activate transcription of the *lacZ* gene, and result in yeast forming blue colonies on X-gal-treated filters. Alternatively, the sequence from the cDNA library introduced into the microorganism may, as was observed in Example 1, include a sequence encoding an activator domain and thus not utilize the acidic activator domain of the yeast GAL4 transcription factor "GAL4-ACT".

Recombinant plasmids from such "blue yeast" are then isolated and transformed back into reporter strains that contain either a wild-type or mutant version of target regulatory sequence fused to the *lacZ* gene. The plasmids that are desired are those that turn the former strains blue, but not the later, indicating that the cloned DNA binding domain is specific for the target regulatory sequence.

Based on presence of an AP2 binding domain in CBF1, CBF2 and CBF3, Applicants believe that an AP2 binding domain is present in a significant number of the environmental stress tolerance regulatory binding proteins. Accordingly, it is believed that the specificity of the above method for screening for gene sequences encoding a regulatory binding protein can optionally be improved by first selecting cDNA from a plant genome library which includes a potential AP2 domain site. This can be routinely done by selecting probes for selecting sequences in the library which include potential AP2 domain sequences.

4. Screening For Expression Of Environmental Stress Tolerance Regulatory Protein

Once one or more microorganisms are selected which are believed to express a protein capable of binding to the target regulatory element and activate expression of the reporter gene, further analysis can be performed to identify and isolate full length cDNAs; i.e. cDNAs that encode the entire protein that binds to the target regulatory sequence. The coding sequence for the protein can then cloned into an expression vector, such as the pET bacterial expression vectors (Novagen), and used to produce the protein at high levels. The protein can then be analyzed by gel retardation experiments (See Example 1F) to confirm that it binds specifically to the target regulatory sequence.

Potential sequences can be further screened using known regulatory gene sequences, such as CBF1, CB2, and CBF3, or the presence of an AP2 domain which is believed to be common to a significant class of this genes. Once identified, particular sequences can be transformed into yeast to test for activation of expression of a reporter gene, for example as described in Example 1E.

5. Screening For Binding To Target Regulatory Sequence

Once a regulatory gene sequence is identified, the sequence can be introduced into a microorganism in order to express the protein encoded by the sequence. A gel shift assay, such as the one described in Example 1F, can then be used to test for *in vitro* binding of the expressed protein to the target DNA regulatory sequence.

Mutagenesis of the target DNA regulatory sequence can also be performed in order to evaluate the binding selectivity of the expressed protein. It is preferred that the expressed protein selectively bind to the target DNA regulatory sequence over related sequences with one or more base differences from the target DNA regulatory sequence. For example, Figure 5 is a photograph of a gel from a shift assay in which CBF1 was shown to selectively bind to the wild-type C-repeat/DRE CCGAC.

6. Altering The Environmental Stress Tolerance of a Plant.

The present invention also provides a method for recombinant engineered plants with a new or altered response to one or more environmental stresses.

According to one embodiment, a copy of a gene native to a plant which encodes a binding protein according to the present invention is recombinantly introduced into the plant such that the plant expresses a recombinant binding protein encoded by the recombinant copy of the gene.

According to another embodiment, a non-native gene which encodes a binding protein according to the present invention is recombinantly introduced into a plant such that the plant expresses a recombinant binding protein encoded by the recombinant non-native gene.

According to yet another embodiment, a native or non-native DNA regulatory sequence is recombinantly introduced into a plant such that the recombinant DNA regulatory sequence regulates the expression of one or more environmental stress tolerance genes in the plant. The plant includes a gene which encodes a binding protein capable of binding to the recombinant DNA regulatory sequence.

In yet another embodiment, a native or non-native promoter is recombinantly introduced into a plant such that the recombinant promoter regulates the expression of a binding protein which binds to a DNA regulatory sequence.

According to each of the above embodiments, unless otherwise specified, the gene encoding the binding protein, the promoter promoting the expression of the binding protein, the DNA regulatory sequence, and the environmental stress tolerance genes may be non-recombinant or recombinant sequences. The recombinant sequences may be native to the plant or may be non-native to the plant. All the above permutations are intended to fall within the scope of the present invention.

As an example, many plants increase in freezing tolerance in response to low non-freezing temperatures, a process known as cold acclimation. A large number of biochemical changes occur during cold acclimation including the activation of *COR* (COld Regulated) genes. These genes, which are also expressed in response to dehydration (e.g., drought and high salinity), are thought to help protect plant cells against the potentially deleterious effects of dehydration associated with freezing.

drought and high salinity stress. Indeed, expression of the *COR15a* gene in plants grown at normal temperatures (22°C) enhances the freezing tolerance of chloroplasts.

By manipulating the expression of *COR* genes, the stress tolerance of crop and horticultural plants could be improved, e.g., engineer broader climate ranges; target stress resistance to stress-sensitive parts of plants; render plants stress-resistant when a stress condition (frost and drought) is imminent. To bring about these effects, however, the expression of the *COR* genes must be manipulated. The gene, *CBF1*, that encodes the transcription factor that binds to the C-repeat/DRE regulatory element present in the promoters of all *COR* genes described to date has been isolated. *CBF1* in yeast activates expression of reporter genes that have been fused to the C-repeat/DRE element. Further, expression of *CBF1* in plants has been shown to activate the expression of *COR* genes.

By introducing modified versions of sequences encoding regulatory binding proteins, such as *CBF1*, into plants, the expression of *COR* genes can be modified, and thereby enhance the freezing and dehydration tolerance of plants.

In each of the above embodiments, expression of the recombinant copy of the regulatory gene may be under the control of the promoter. The promoter may increase the level at which the regulatory gene is expressed, express the regulatory gene without being induced by an environmental stress or express the regulatory gene in response to a different form or degree of environmental stress that would otherwise be needed to induce expression of the regulatory gene. For example, a strong constitutive promoter could be used to cause increased levels of *COR* gene expression in both non-stress and stressed plants which in turn, results in enhanced freezing and dehydration tolerance. A tissue specific promoter could be used to alter COR gene expression in tissues that are highly sensitive to stress (and thereby enhance the stress tolerance of these tissues). A promoter could be used which turns on at a temperature that is warmer than the temperature at which the plant normally exhibits cold tolerance. This would enable the cold tolerance thermostat of a plant to be altered. Similarly, a promoter can be used which turns on at a dehydration condition that is wetter than the dehydration condition at which the plant normally exhibits dehydration tolerance. This would enable the level at which a plant

responds to dehydration to be altered. An inducible promoter could be used such that gene expression could be induced by application of an exogenous inducer (e.g., induce *COR* genes when a frost is imminent). This would enable stress tolerance to be induced externally by a grower whenever desired.

It is envisioned that the present invention can be used to introduce, change and/or augment the environmental stress tolerance of a plant by introducing and causing the expression of environmental stress tolerance in a manner which the plant does not exhibit in its native form. For example, by using different promoters in combination with recombinant regulatory genes, native environmental stress tolerance genes can be expressed independent of environmental stress, made responsive to different levels or types of environmental stress, or rendered inducible independent of an environmental stress.

By recombinantly introducing a native environmental stress tolerance gene into a plant in combination with a recombinant regulatory gene under the control of an inducible promoter, a plant can be engineered which includes its native environmental stress tolerance as well as inducible environmental stress tolerance. This might be useful for inducing a cold stress tolerance reaction in anticipation of a frost.

By recombinantly introducing a non-native environmental stress tolerance gene into a plant in combination with a recombinant regulatory gene, a plant can be engineered which includes environmental stress tolerance properties that the plant would not otherwise have. In this regard, plants from warmer climates can be engineered to include one or more cold tolerance genes along with a regulatory gene needed to cause expression of the cold tolerance genes in the plant so that the engineered plant can survive better in a colder climate. Similarly, a plant can be engineered to include one or more dehydration tolerance genes along with a regulatory gene needed to cause expression of the dehydration tolerance gene so that the engineered plant can grow better in a dryer climate. In this regard, it should be possible to take a plant which grows well in a first climate and engineer it to include stress tolerance genes and regulatory genes native to a second climate so that the plant can grow well in the second climate.

By modifying the promoter controlling the expression of the gene encoding a binding protein which regulates the expression of environmental stress tolerance genes, the operation of native, non-recombinant environmental stress tolerance genes and regulatory genes can be changed. For example, the conditions under which the stress tolerance genes are expressed can be changed. Expression can also be rendered inducible by an exogenous agent.

7. Methods For Detecting Stress Tolerance Regulatory Gene Homologs.

Once one DNA sequence encoding an environmental stress tolerance regulatory binding protein has been identified, several methods are available for using that sequence and knowledge about the protein it encodes to identify homologs of that sequence from the same plant or different plant species. For example, let us assume that a cDNA encoding a first target binding domain has been isolated from plant species "A." The DNA sequence encoding the first target DNA regulatory sequence could be radiolabeled and used to screen cDNA libraries of plant species "A," or any other plant species, for DNA inserts that encode proteins related to the first target DNA regulatory sequence. This could be done by screening colony or phage "lifts" using either high (Tm of about -10°C) or low (Tm of about -30°C or lower) stringency DNA hybridization conditions (Sambrook, J. et al, Molecular Cloning. A Laboratory Manual Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY., 2nd Ed. (1989)). cDNA inserts that hybridize with the first target DNA regulatory sequence could be sequenced and compared to the original first target DNA regulatory sequence. If the insert is confirmed to encode a polypeptide similar to the first target DNA regulatory sequence, the insert could be cloned into an expression vector to produce the encoded protein. The protein would then be analyzed by gel retardation experiments to confirm that it binds specifically to the first target DNA regulatory sequence. It is recognized that not all proteins that bind to a first target DNA regulatory sequence will be transcriptional activators. However, a number of routine tests may be performed in order to determine whether a particular protein is in fact a transcriptional activator. One test involves expressing the protein in yeast strains which contain the target DNA regulatory sequence fused

to the *lacZ* reporter gene, as described above. If the protein is a transcriptional activator, it should activate expression of the reporter gene and result in blue colonies.

Another test is a plant transient assay. In this case, a reporter gene, such as GUS, carrying the target DNA regulatory sequence as an upstream activator is introduced into plant cells (e.g. by particle bombardment) with or without a the putative transcriptional activator under control of a constitutive promoter. If the protein is an activator, it will stimulate expression of the reporter (this may be further enhanced if the plant material is placed at low temperature or is subjected to water stress as the C-repeat/DRE is responsive to low temperature and dehydration).

Significantly, once a target DNA regulatory sequence is identified, the sequence can be fused to any potential activator or repressor sequence to modify expression of plant genes that carry the target regulatory sequence as a control element. That is, the DNA regulatory sequence can be used to target "managed" expression of the battery of environmental stress tolerance related genes in a given plant species.

It is possible that the target DNA regulatory sequence of the regulatory element that imparts environmental stress tolerance related gene expression in plant species "A" might be slightly different from the analogous target DNA regulatory element that imparts environmental stress tolerance in species "B." Thus, optimal regulation of the battery of environmental stress tolerance related genes in a given species may require the use of the regulatory binding proteins from that or a closely related plant species. Knowledge of gene sequences which encode for proteins which bind to the DNA regulatory sequence of the regulatory element, in combination with knowledge of the DNA regulatory sequence, greatly simplify the identification of sequences encoding binding proteins native to the plant species.

With the advent of fast and efficient DNA sequencing technologies, the number of plant genomes recorded on computer databases is growing rapidly. These computer databases can be used to search for homologs to CBF1, CBF2, CBF3 as well as other sequences which encode binding proteins which regulate cold tolerance genes. As more and more binding protein sequences are identified and the number of computerized plant genome databases increase, searching computer

databases for additional sequences encoding binding proteins which regulate cold tolerance genes will become increasingly simplified.

EXAMPLES

1. Isolation and Analysis of *Arabidopsis Thaliana* cDNA Clone (CBF1) Encoding C-repeat/DRE Binding Factor

The following example describes the isolation of an *Arabidopsis thaliana* cDNA clone that encodes a C-repeat/DRE binding factor, CBF1 (C-repeat/DRE Binding Factor 1). Expression of CBF1 in yeast was found to activate transcription of reporter genes containing the C-repeat/DRE (CCGAC) as an upstream activator sequence. Meanwhile, CBF1 did not activate transcription of mutant versions of the CCGAC binding element, indicating that CBF1 is a transcription factor that binds to the C-repeat/DRE. Binding of CBF1 to the C-repeat/DRE was also demonstrated in gel shift assays using recombinant CBF1 protein expressed in *Escherichia coli*. Analysis of the deduced CBF1 amino acid sequence indicated that the protein has a potential nuclear localization sequence, a possible acidic activation domain and an AP2 domain, a DNA-binding motif of about 60 amino acids that is similar to those present in Arabidopsis proteins APETALA2, AINTEGUMENTA and TINY, the tobacco ethylene response element binding proteins, and numerous other plant proteins of unknown function.

A. Materials

Plant material and cold treatment. *A thaliana* (L.) Heyn. ecotype RLD plants were grown in pots in controlled environment chambers at 22°C under constant illumination with cool-white fluorescent lamps (~100 μmol $m^{-2}s^{-1}$) essentially as described (Gilmour, S. J., Plant Physiol. **87**:745-750 (1988)). Plants were cold-treated by placing pots in a cold room at 2.5°C under constant illumination with cool-white florescent lamps (~25 μmol $m^{-2}s^{-1}$) for the indicated times.

Arabidopsis cDNA expression library. The Arabidopsis pACT cDNA expression library was constructed by John Walker and colleagues

(NSF/DOE/USDA Collaborative Research in Plant Biology Program grant USDA 92-37105-7675) and deposited in the Arabidopsis Biological Resource Center (stock #CD4-10).

Yeast reporter strains. Oligonucleotides (Table 1) (synthesized at the MSU Macromolecular Structure Facility) encoding either wild-type or mutant versions of the C-repeat/DRE were ligated into the *Bgl*II site of the *lacZ* reporter vector pBgl-lacZ (Li, J. J. and I. Herskowitz, Science **262**:1870-1874 (1993); kindly provided by Joachim Li). The resulting reported constructs were integrated into the *ura3* locus of *Saccharomyces cerevisiae* strain GGY1 (*MATα Δgal4 Δgal80 ura3 leu2 his3 ade2 tyr*) (Li, J. J. and I. Herskowitz, Science **262**:1870-1874 (1993); provided by Joachim Li) by transformation and selection for uracil prototrophy.

E. coli strains. *Escherichia coli* strain GM2163 containing plasmid pEJS251 was deposited under the Budapest Treaty on May 17, 1996 with the American Type Culture Collection, Rockville, Maryland as ATCC 98063. It is available by name and number pursuant to the provisions of the Budapest Treaty.

TABLE 1

Oligonucleotides encoding wild type and mutant versions of the C-repeat/DRE

Oligonucleotide	C-repeat /DRE*	Sequence	SEQ ID NO:
MT50	*COR15a*	gatcATTTCATGG**CCGAC**CTGCTTTTT	3
MT52	M1*COR15a*	CACAA**TTT**CAa**G**aattca**C**TGCTTTTTT	4
MT80	M2*COR15a*	gatcATTTCATGG**t**at**g**tCTGCTTTTT	5
MT125	M3*COR15a*	gatcATTTCATGG**a**atca**C**TGCTTTTT	6
MT68	*COR15a*	gatcACTTGATGG**CCGAC**CTCTTTTTT	7
MT66	*COR78-1*	gatcAATATACTA**CCGAC**ATGAGTTCT	8
MT86	*COR78-2*	ACTA**CCGAC**ATGAGTTCCAAAAAGC	9

*The C-repeat/DRE sequences tested are either wild-type found in the promoters of *COR15a* (Baker, S. S., et al., Plant. mol. Biol. **24**:701-713 (1994)), *COR15b* or *COR78/RD29A* (Horvath, D. P., et al., Plant Physiol. **103**:1047-1053 (1993); Yamaguchi-shinozaki, K., et al., The Plant Cell **6**:251-264 (1994)) or are mutant versions of the *COR15a* C-repeat/DRE (M1*COR15a*, M2*COR15a* and M3*COR15a*).

#Uppercase letters designate bases in wild type C-repeat/DRE sequences. The core CCGAC sequence common to the above sequences is indicated in bold type. Lowercase letters at the beginning of a sequence indicate bases added to facilitate cloning. The lowercase letters that are underlined indicate the mutations in the C-repeat/DRE sequence of *COR15a*.

B. **Methods**

Screen of Arabidopsis cDNA library. The Arabidopsis pACT cDNA expression library was screened for clones encoding C-repeat/DRE binding domains by the following method. The cDNA library, harbored in *Escherichia coli* BNN132, was amplified by inoculating 0.5 ml of the provided glycerol stock into 1 L of M9 minimal glucose medium (Sambrook, J. et al, Molecular Cloning. A Laboratory Manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY., 2nd Ed. (1989)) and shaking the bacteria for 20 h at 37°C. Plasmid DNA was isolated and purified by cesium chloride density gradient centrifugation (Sambrook et al (1989)) and transformed into the yeast GGY1 reporter strains selecting for leucine prototrophy. Yeast transformants that had been grown for 2 or 3 days at 30°C were overlaid with either a nitrocellulose membrane filter (Schleicher and Schuell, Keene, NH) or Whatman #50 filter paper (Hillsboro, OR) and incubated overnight at 30°C. The yeast impregnated filters were then lifted from the plate and treated with X-gal (5-bromo-4-chloro-3-indolyl β-D-galactosidase) to assay colonies for β-galactosidase activity (Li, J. J. and I. Herskowitz, Science **262**:1870-1874 (1993)). Plasmid DNA from "positive" transformants (those forming blue colonies on the X-gal-treated filters) was recovered (Strathern, J. N., and D. R. Higgens, Methods Enzymol. **194**:319-329 (1991)), propagated in *E. coli* DH5α and transformed back into the yeast reporter strains to confirm activity.

Yeast transformation and quantitative β-galactosidase assays. Yeast were transformed by either electroporation (Becker, D. M., et al., Methods Enzymol. **194**:182-187 (1991)) or the lithium acetate/carrier DNA method (Schiestl, R. H., et al., Current Genetics **16**:339-346 (1989)). Quantitative *in vitro* β-galactosidase assays were done as described (Rose, M., et al., Methods Enzymol. **101**:167-180 (1983)).

Expression of CBF1 protein in *E. coli* and yeast. CBF1 was expressed in *E. coli* using the pET-28a(+) vector (Novagen, Madison, WI). The *Bgl*II-*Bcl*I restriction fragment of pACT-11 encoding CBF1 was ligated into the

*Bam*HI site of the vector bringing CBF1 under control of the T7 phage promoter. The construct resulted in a "histidine tag," a thrombin recognition sequence and a "T7 epitope tag" being fused to the amino terminus of CBF1. The construct was transformed into *E. coli* BL21 (DE3) and the recombinant CBF1 protein was expressed as recommended by the supplier (Novagen). Expression of CBF1 in yeast was accomplished by ligating restriction fragments encoding CBF1 (the *Bcl*I-*Bgl*II and *Bgl*II-*Bgl*II fragments from pACT-11) into the *Bgl*II site of pDB20.1 (Berger, S. L., et al., Cell **70**:251-265 (1992); kindly provided by Steve Triezenberg) bringing CBF1 under control of the constitutive *ADC1* (alcohol dehydrogenase constitutive 1) promoter.

Gel shift assays. The presence of expressed protein which binds to a C-repeat/DRE binding domain was evaluated using the following gel shift assay. Total soluble *E. coli* protein (40 ng) was incubated at room temperature in 10 µl of 1X binding buffer [15 mM HEPES (pH 7.9), 1 mM EDTA, 30mM KCl, 5% glycerol, 5% BSA, 1mM DTT) plus 50 ng poly(dI-dC):poly(dI-dC) (Pharmacia, Piscataway, NJ) with or without 100 ng competitor DNA. After 10 min, probe DNA (1 ng) that was ^{32}P-labeled by end-filling (Sambrook et al, 1989) was added and the mixture incubated for an additional 10 min. Samples were loaded onto polyacrylamide gels (4% w/v) and fractionated by electrophoresis at 150V for 2h (Sambrook et al). Probes and competitor DNAs were prepared from oligonucleotide inserts ligated into the *Bam*HI site of pUC118 (Vieira, J., et al., Methods Enzymol. **153**:3-11 (1987)). Orientation and concatenation number of the inserts were determined by dideoxy DNA sequence analysis (Sambrook, et al, (1989)). Inserts were recovered after restriction digestion with *Eco*RI and *Hind*III and fractionation on polyacrylamide gels (12% w/v) (Sambrook et al, 1989).

Northern and southern analysis. Northern and southern analysis was performed as follows. Total RNA was isolated from Arabidopsis (Gilmour, S. J., et al., Plant Physiol. **87**:745-750 (1988)) and the poly(A)$^+$ fraction purified using oligo dT cellulose (Sambrook, et al (1989)). Northern transfers were prepared and hybridized as described (Hajela, R. K., et al., Plant Physiol. **93**:1246-1252 (1990))

except that high stringency wash conditions were at 50°C in 0.1X SSPE [X SSPE is 3.6 M NaCl, 20 mM EDTA, 0.2 M Na_2-HPO_4 (pH7.7)], 0.5% SDS. Membranes were stripped in 0.1 X SSPE, 0.5% SDS at 95°C for 15 min prior to re-probing. Total Arabidopsis genomic DNA was isolated (Stockinger, E. J., et al., J. Heredity, **87**:214-218 (1996)) and southern transfers prepared (Sambrook et al 1989) using nylon membranes (MSI, Westborough, MA). High stringency hybridization and wash conditions were as described by Walling et al (Walling, L. L., et al., Nucleic Acids Res. **16**:10477-10492 (1988)). Low stringency hybridization was in 6X SSPE, 0.5% SDS, 0.25% low fat dried milk at 60°C. Low stringency washes were in 1X SSPE, 0.5% SDS at 50°C. Probes used for the entire CBF1 coding sequence and 3' end of CBF1 were the *Bcl*I/*Bgl*II and *Eco*RV/*Bgl*II restriction fragments from pACT-11, respectively, that had been gel purified (Sambrook et al (1989)). DNA probes were radiolabeled with ^{32}P-nucleotides by random priming (Sambrook). Autoradiography was performed using hyperfilm-MP (Amersham, Arlington Heights, IL). Radioactivity was quantified using a Betascope 603 blot analyzer (Betagen Corp., Waltham, MA).

C. Screen of Arabidopsis cDNA library for sequence encoding a C-repeat/DRE binding domain.

The "one-hybrid" strategy (Li, J. J. and I. Herskowitz, Science **262**:1870-1874 (1993)) was used to screen for Arabidopsis cDNA clones encoding a C-repeat/DRE binding domain. In brief, yeast strains were constructed that contained a *lacZ* reporter gene with either wild-type or mutant C-repeat/DRE sequences in place of the normal UAS (upstream activator sequence) of the *GAL1* promoter.

Figures 1A and 1B show how the yeast reporter strains were constructed. Figure 1A is a schematic diagram showing the screening strategy. Yeast reporter strains were constructed that carried C-repeat/DRE sequences as UAS elements fused upstream of a *lacZ* reporter gene with a minimal *GAL1* promoter. The strains were transformed with an Arabidopsis expression library that contained random cDNA inserts fused to the GAL4 activation domain (GAL4-ACT) and screened for blue colony formation on X-gal-treated filters. Figure 1B is a chart showing activity of the "positive" cDNA clones in yeast reporter strains. The oligonucleotides (oligos)

used to make the UAS elements, and their number and direction of insertion, are indicated by the arrows.

Yeast strains carrying these reporter constructs produced low levels of β-galactosidase and formed white colonies on filters containing X-gal. The reporter strains carrying the wild-type C-repeat/DRE sequences were transformed with a DNA expression library that contained random Arabidopsis cDNA inserts fused to the acidic activator domain of the yeast GAL4 transcription factor, "GAL4-ACT" (Figure 1A). The notion was that some of the clones might contain a cDNA insert encoding a C-repeat/DRE binding domain fused to GLA4-ACT and that such a hybrid protein could potentially bind upstream of the *lacZ* reporter genes carrying the wild type C-repeat/DRE sequence, activate transcription of the *lacZ* gene and result in yeast forming blue colonies on X-gal-treated filters.

Upon screening about 2 x 10^6 yeast transformants, three "positive" cDNA clones were isolated; i.e., clones that caused yeast strains carrying *lacZ* reporters fused to wild-type C-repeat/DRE inserts to form blue colonies on X-gal-treated filters (Figure 1B). The three cDNA clones did not cause a yeast strain carrying a mutant C-repeat/DRE fused to *LacZ* to turn blue (Figure 1B). Thus, activation of the reporter genes by the cDNA clones appeared to be dependent on the C-repeat/DRE sequence. Restriction enzyme analysis and DNA sequencing indicated that the three cDNA clones had an identical 1.8 kb insert (Figure 2A). One of the clones, designated pACT-11, was chosen for further study.

D. Identification of 24 kDa polypeptide with an AP2 domain encoded by pACT-11.

Figures 2A, 2B, 2C and 2D provide an analysis of the pACT-11 cDNA clone. Figure 2A is a schematic drawing of the pACT-11 cDNA insert indicating the location and 5' to 3' orientation of the 24 kDa polypeptide and 25s rRNA sequences. The cDNA insert was cloned into the *Xho*I site of the pACT vector. Figure 2B is a DNA and amino acid sequence of the 24 kDa polypeptide (SEQ ID NO:1 and SEQ ID NO:2). The AP2 domain is indicated by a double underline. The basic amino acids that potentially act as a nuclear localization signal are indicated with asterisks. The *Bcl*I site immediately upstream of the 24 kDa polypeptide used in subcloning the 24

kDa polypeptide and the *Eco*RV site used in subcloning the 3' end of CBF1 are indicated by single underlines. Figure 2C is a schematic drawing indicating the relative positions of the potential nuclear localization signal (NLS), the AP2 domain and the acidic region of the 24 kDa polypeptide. Numbers indicate amino acid residues. Figure 2D is a chart showing comparison of the AP2 domain of the 24 kDa polypeptide with that of the tobacco DNA binding protein EREBP2 (Okme-Takagi, M., et al., The Plant Cell **7**:173-182 (1995) SEQ ID NOS: 10 and 11). Identical amino acids are indicated with single lines; similar amino acids are indicated by double dots; amino acids that are invariant in AP2 domains are indicated with asterisks (Klucher, K. M., et al., The Plant Cell **8**:137-153 (1996)); and the histidine residues present in CBF1 and TINY (Wilson, K., et al., The Plant Cell **8**:659-671 (1996)) that are tyrosine residues in all other described AP2 domains are indicated with a caret. A single amino acid gap in the CBF1 sequence is indicated by a single dot.

Our expectation was that the cDNA insert in pACT-11 would have a C-repeat/DRE binding domain fused to the yeast GAL4-ACT sequence. However, DNA sequence analysis indicated that an open reading frame of only nine amino acids had been added to the C-terminus of GAL4-ACT. It seemed highly unlikely that such a short amino acid sequence could comprise a DNA binding domain. Also surprising was the fact that about half of the cDNA insert in pACT-11 corresponded to 25s rRNA sequences (Figure 2A). Further analysis, however, indicated that the insert had an open reading frame, in opposite orientation to the GAL4-ACT sequence, deduced to encode a 24 kDa polypeptide (Figure 2A-C). The polypeptide has a basic region that could potentially serve as a nuclear localization signal (Raikhel, N., Plant Physiol. **100**:1627-1632 (1992)) and an acidic C-terminal half (pI of 3.6) that could potentially act as an acidic transcription activator domain (Hahn, S., Cell **72**:481-483 (1993)). A search of the nucleic acid and protein sequence databases indicated that there was no previously described homology of the 24 kDa polypeptide. However, the polypeptide did have an AP2 domain (Jofuku, K. D., et al., The Plant Cell **6**:1211-1225 (1994)) (Figures 2B, D), a DNA binding motif of about 60 amino acids (Ohme-Takagi, M., et al., The Plant Cell **7**:173-182 (1994)) that is present in numerous plant proteins including the APETALA2 (Jofuku, K. D., et

al., The Plant Cell **6**:1211-1225 (1994)), AINTEGUMENTA (Klucher, K. M., et al., The Plant Cell **8**:137-153 (1996); Elliot, R. C., et al., The Plant Cell **8**:155-168 (1996)) and TINY (Wilson, K., et al., The Plant Cell **8**:659-671 (1996)) proteins of Arabidopsis and the EREBPs (ethylene response element binding proteins) of tobacco (Ohme-Takagi, M., et al., The Plant Cell **7**:173-182 (1995)).

E. 24 kDa polypeptide binds to the C-repeat/DRE and activates transcription in yeast.

We hypothesized that the 24 kDa polypeptide was responsible for activating the *lacZ* reporter genes in yeast. To test this, the *Bcl*I-*Bgl*II fragment of pACT-11 containing the 24 kDa polypeptide, and the *Bgl*II-*Bgl*II fragment containing the 24 kDa polypeptide plus a small portion of the 25s rRNA sequence, was inserted into the yeast expression vector pDB20.1

Figure 3 is a chart showing activation of reporter genes by the 24 kDa polypeptide. Restriction fragments of pACT-11 carrying the 24 kDa polypeptide (*Bcl*I-*Bgl*II) or the 24 kDa polypeptide plus a small amount of 25s RNA sequence (*Bgl*II-*Bgl*II) were inserted in both orientations into the yeast expression vector pDB20.1 (see Figure 2A and 2B for location of *Bcl*I and *Bgl*II restriction sites). These "expression constructs" were transformed into yeast strains carrying the *lacZ* reporter gene fused to direct repeat dimers of either the wild-type *COR15a* C-repeat/DRE (oligonucleotide MT50) or the mutant M2*COR15a* C-repeat/DRE (oligonucleotide MT80). The specific activity of β-galactosidase (nmoles o-nitrophenol produced/min^{-1} x mg protein^{-1}) was determined from cultures grown in triplicate. Standard deviations are indicated. Abbreviations: pADC1, *ADC*1 promoter; tADC1, *ADC1* terminator.

Plasmids containing either insert in the same orientation as the *ADC1* promoter stimulated synthesis of β-galactosidase when transformed into yeast strains carrying the *lacZ* reporter gene fused to a wild-type *COR15a* C-repeat/DRE (Figure 3). The plasmids did not, however, stimulate synthesis of β-galactosidase when transformed into yeast strains carrying *lacZ* fused to a mutant version of the *COR15a* C-repeat/DRE (Figure 3). These data indicated that the 24 kDa polypeptide could bind to the wild-type C-repeat/DRE and activate expression for the

lacZ reporter gene in yeast. Additional experiments indicated that the 24 kDa polypeptide could activate expression of the *lacZ* reporter gene fused to either a wild-type *COR78* C-repeat/DRE (dimer of MT66) or a wild-type *COR15b* C-repeat/DRE (dimer of MT 68) (not shown). A plasmid containing the *Bcl*I-*Bgl*II fragment (which encodes only the 24 kDa polypeptide) cloned in opposite orientation to the *ADC1* promoter did not stimulate synthesis of β-galactosidase in reporter strains carrying the wild-type *COR15a* C-repeat/DRE fused to *lacZ* (Figure 3). In contrast, a plasmid carrying the *Bgl*II-*Bgl*II fragment (containing the 24 kDa polypeptide plus some 25s rRNA sequences) cloned in opposite orientation to the *ADC1* promoter produced significant levels of β-galactosidase in reporter strains carrying the wild-type *COR15a* C-repeat/DRE (Figure 3). Thus, a sequence located closely upstream of the 24 kDa polypeptide was able to serve as a cryptic promoter in yeast, a result that offered an explanation for how the 24 kDa polypeptide was expressed in the original pACT-11 clone.

F. Gel shift analysis indicates that the 24 kDa polypeptide binds to the C-repeat/DRE.

Gel shift experiments were conducted to demonstrate further that the 24 kDa polypeptide bound to the C-repeat/DRE. Specifically, the open reading frame for the 24 kDa polypeptide was inserted into the pET-28a(+) bacterial expression vector (see Materials and Methods) and the resulting 28 kDa fusion protein was expressed at high levels in *E. coli*. (Figure 4).

Figure 4 is a photograph of an electrophoresis gel showing expression of the recombinant 24 kDa polypeptide in *E. coli*. Shown are the results of SDS-PAGE analysis of protein extracts prepared from *E. coli* harboring either the expression vector alone (vector) or the vector plus an insert encoding the 24 kDa polypeptide in sense (sense insert) or antisense (antisense insert) orientation. The 28 kDa fusion protein (see Materials and Methods) is indicated by an arrow.

Figure 5 is a photograph of a gel for shift assays indicating that CBF1 binds to the C-repeat/DRE. The C-repeat/DRE probe (1 ng) used in all reactions was a ^{32}P-labeled dimer of the oligonucleotide MT50 (wild type C-repeat/DRE from *COR15a*). The protein extracts used in the first four lanes were either bovine serum albumin

(BSA) or the indicated CBF1 sense, antisense and vector extracts described in Figure 4. The eight lanes on the right side of the figure used the CBF1 sense protein extract plus the indicated competitor C-repeat/DRE sequences (100 ng). The numbers 1X, 2X and 3X indicate whether the oligonucleotides were monomers, dimers or trimers, respectively, of the indicated C-repeat/DRE sequences.

Protein extracts prepared from *E. coli* expressing the recombinant protein produced a gel shift when a wild-type *COR15a* C-repeat/DRE was used as probe (Figure 5). No shift was detected with BSA or *E. coli* extracts prepared from strains harboring the vector alone, or the vector with an antisense insert for the 24 kDa polypeptide. Oligonucleotides encoding wild-type C-repeat/DRE sequences from *COR15a* or *COR78* competed effectively for binding to the *COR15a* C-repeat/DRE probe, but mutant version of the *COR15a* C-repeat/DRE did not (Figure 5). These *in vitro* results corroborated the *in vivo* yeast expression studies indicating that the 24 kDa polypeptide binds to the C-repeat/DRE sequence. The 24 kDa polypeptide was thus designated CBF1 (C-repeat/DRE binding factor 1) and the gene encoding it named *CBF1*.

G. CBF1 is a unique or low copy number gene.

Figure 6 is a photograph of a southern blot analysis indicating *CBF1* is a unique or low copy number gene. Arabidopsis DNA (~1µg) was digested with the indicated restriction endonucleases and southern transfers were prepared and hybridized with a ^{32}P-labeled probe encoding the entire CBF1 polypeptide.

The hybridization patterns observed in southern analysis of Arabidopsis DNA using the entire *CBF1* gene as probe were relatively simple indicating that *CBF1* is either a unique or low copy number gene (Figure 6). The hybridization patterns obtained were not altered if only the 3' end of the gene was used as the probe (the *Eco*RV/*Bgl*II restriction fragment from pACT-11 encoding the acidic region of CBF1, but not the AP2 domain) or if hybridization was carried out at low stringency (not shown).

H. CBF1 transcript level response to low temperature.

Figures 7A, 7B and 7C relate to *CBF1* transcripts in control and cold-treated Arabidopsis. Figure 7A is a photograph of a membrane RNA isolated from Arabidopsis plants that were grown at 22°C or grown at 22°C and transferred to 2.5°C for the indicated times. Figures 7B and 7C are graphs showing relative transcript levels of *CBF1* and *COR15a* in control and cold-treated plants. The radioactivity present in the samples described in Figure 7A were quantified using a Betascope 603 blot analyzer and plotted as relative transcript levels (the values for the 22°C grown plants being arbitrarily set as 1) after adjusting for differences in loading using the values obtained with the pHH25 probe.

Based on Figures 7A-7C, northern analysis indicated that the level of *CBF1* transcripts increased about 2 to 3 fold in response to low temperature (Figure 7B). In contrast, the transcript levels for *COR15a* increased approximately 35 fold in cold-treated plants (Figure 7C). Only a singly hybridizing band was observed for *CBF1* at either high or low stringency with probes for either the entire *CBF1* coding sequence or the 3' end of the gene (the *Eco*RV/*Bgl*II fragment of pACT-11) (not shown). The size of the *CBF1* transcripts was about 1.0 kb.

I. Discussion Of Experimental Results.

The above example regarding CBF1 represents the first identification of a gene sequence which encodes a protein capable of binding to the C-repeat/DRE sequence CCGAC. The experimental results presented evidence that CBF1 binds to the C-repeat/DRE both *in vitro* via gel shift assays and *in vivo* via yeast expression assays. Further, the results demonstrate that CBF1 can activate transcription of reporter genes in yeast that contain the C-repeat/DRE.

The results of the southern analysis indicate that *CBF1* is a unique or low copy number gene in Arabidopsis. However, the CBF1 protein contains a 60 amino acid motif, the AP2 domain, that is evolutionary conserved in plants (Weigel, D., The plant Cell **7**:388-389 (1995)). It is present in the APETALA2 (Jofuku, K. D., et al., The Plant Cell **6**:1211-1225 (1994)), AINTEGUMENTA (Klucher, K. M., et al., the Plant Cell **8**:137-153 (1996; and Elliot, R. C., et al., The Plant Cell **8**:155-168 (1996)), TINY (Wilson, K., et al., The Plant Cell **8**:659-671 (1996)) and cadmium-

induced (Choi, S.-Y., et al., Plant Physiol. **108**:849 (1995)) proteins of Arabidopsis and the EREBPs of tobacco (Ohme-Takagi, M. et al., The Plant Cell **7**:173-182 (1995)). In addition, a search of the GenBank expressed sequence tagged cDNA database indicates that there is one cDNA from *B. napus*, two from *Ricinus communis*, and more than 25 from Arabidopsis and 15 from rice, that are deduced to encode proteins with AP2 domains. The results of Ohme-Takagi and Shinshi (Ohme-Takagi, M., et al., The Plant Cell **7**:173-182 (1995)) indicate that the function of the AP2 domain is DNA-binding; this region of the putative tobacco transcription factor EREBP2 is responsible for its binding to the *cis*-acting ethylene response element referred to as the GCC-repeat. As discussed by Ohme-Takagi and Shinshi (Ohme-Takagi, M., et al., the Plant Cell **7**:173-182 (1995)), the DNA-binding domain of EREBP2 (the AP2 domain) contains no significant amino acid sequence similarities or obvious structural similarities with other known transcription factors or DNA binding motifs. Thus, the domain appears to be a novel DNA-binding motif that to date, has only been found in plant proteins.

It is believed that the binding of CBF1 to the C-repeat/DRE involves the AP2 domain. In this regard, it is germane to note that the tobacco ethylene response element, AGCCGCC, closely resembles the C-repeat/DRE sequences present in the promoters of the Arabidopsis genes *COR15a*, GGCCGAC, and *COR78/RD29A*, TACCGAC. Applicants believe that CBF1, the EREBPs and other AP2 domain proteins are members of a superfamily of DNA binding proteins that recognize a family of *cis*-acting regulatory elements having CCG as a common core sequence. Differences in the sequence surrounding the CCG core element could result in recruitment of different AP2 domain proteins which, in turn, could be integrated into signal transduction pathways activated by different environmental, hormonal and developmental cues. Such a scenario is akin to the situation that exists for the ACGT-family of *cis*-acting elements (Foster et al., FASEB J. **8**:192-200 (1994)). In this case, differences in the sequence surrounding the ACGT core element result in the recruitment of different bZIP transcription factors involved in activating transcription in response to a variety of environmental and developmental signals.

The results of the yeast transformation experiments indicate that CBF1 has a domain that can serve as a transcriptional activator. The most likely candidate for

this domain is the acidic C-terminal half of the polypeptide. Indeed, random acidic amino acid peptides from *E. coli* have been shown to substitute for the GAL4 acidic activator domain of GAL4 in yeast (Ma, J. and M. Ptashne, Cell **51**:113-199 (1987)). Moreover, acidic activator domains have been found to function across kingdoms (Hahn, S., Cell **72**:481-483 (1993)); the yeast GAL4 acidic activator, for instance, can activate transcription in tobacco (Ma, J., et al., Nature **334**:631-633 (1988)). It has also been shown that certain plant transcription factors, such as Vp1 (McCarty, D. R., et al., Cell **66**:895-905 (1991)), have acidic domains that function as transcriptional activators in plants. Significantly, the acidic activation domains of the yeast transcription factors VP16 and GCN4 require the "adaptor" proteins ADA2, ADA3, and GCN5 for full activity (see Guarente, L., Trends Biochem. Sci. **20**:517-521 (1995)). These proteins form a heteromeric complex (Horiuchi, J., et al., Mol. Cell Biol. **15**:1203-1209 (1995)) that bind to the relevant activation domains. The precise mechanism of transcriptional activation is not known, but appears to involve histone acetylation: there is a wealth of evidence showing a positive correlation between histone acetylation and the transcriptional activity of chromatin (Wolffe, A. P., Trends Biochem. Sci. **19**:240-244 (1994)) and recently, the GCN5 protein has been shown to have histone acetyltransferase activity (Brownell, J. E., et al., Cell **84**:843-851 (1996)). Genetic studies indicate that CBF1, like VP16 and GCN4, requires ADA2, ADA3 and GCN5 to function optimally in yeast. The fundamental question thus raised is whether plants have homologs of ADA2, ADA3 and GCN5 and whether these adaptors are required for CBF1 function (and function of other transcription factors with acidic activator regions) in Arabidopsis.

A final point regards regulation of CBF1 activity. The results of the northern analysis indicate that *CBF1* transcript levels increase only slightly in response to low temperature, while those for *COR15a* increase dramatically (Fig. 7). Thus, unlike in yeast, it would appear that transcription of *CBF1* in Arabidopsis at warm temperatures is not sufficient to cause appreciable activation of promoters containing the C-repeat/DRE. The molecular basis for this apparent low temperature activation of CBF1 in Arabidopsis is not known. One intriguing possibility, however is that CBF1 might be modified at low temperature in Arabidopsis resulting in either stabilization of the protein, translocation of the protein

from the cytoplasm to the nucleus, or activation of either the DNA binding domain or activation domain of the protein. Such modification could involve a signal transduction pathway that is activated by low temperature. Indeed, as already discussed, cold-regulated expression of COR genes in Arabidopsis and alfalfa appears to involve a signal transduction pathway that is activated by low temperature-induced calcium flux (Knight, H., et al., The Plant Cell **8**:489-503 (1996); Knight, M. R., et al., Nature **352**:524-526 (1991); Monroy, A. F., et al, Plant Physiol. **102**:1227-1235 (1993); Monroy, A. F., and R. S., The Plant Cell, **7**:321-331 (1995)). It will, therefore, be of interest to determine whether CBF1 is modified at low temperature, perhaps by phosphorylation, and if so, whether this is dependent on calcium-activated signal transduction.

2. Use of CBF1 To Induce Cold Regulated Gene Expression in Nonacclimated *Arabidopsis* Plants.

The following example demonstrates that increased expression of CBF1 induces *COR* gene expression in nonacclimated *Arabidopsis* plants. Transgenic Arabidopsis plants that overexpress *CBF1* were created by placing a cDNA encoding CBF1 under the control of the strong cauliflower mosaic virus (CaMV) 35*S* promoter and transforming the chimeric gene into *Arabidopsis* ecotype RLD plants (Standard procedures were used for plasmid manipulations (J. Sambrook, et al., Molecular Cloning, A Laboratory Manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, ed. 2, (1989)). The CBF1-containing *AseI*-*BglII* fragment from pACT-BgI + (Stockinger, E. J., et al., Proc. Natl. Acad. Sci. U.S.A. **94**:1035 (1997)) was gel-purified, *Bam*HI linkers were ligated to both ends and the fragment was inserted into the *Bam*HI site in pCIB710 (S. Rothstein, et al., Gene **53**:153-161 (1987)) which contains the CaMV 35*S* promoter and terminator. The chimeric plasmid was linearized at the *KpnI* site and inserted into the *KpnI* site of the binary vector pCIB10g (Ciba-Geigy, Research Triangle Park, NC). The plasmid was transformed into *Agrobacterium tumefaciens* strain C58C1 (pMP90) by electroporation. *Arabidopsis* plants were transformed by the vacuum infiltration procedure (N. Bechtold, J. Ellis, and G. Pelletier, C. R. Acad. Sci. Paris, Life Sci. **316**:1194-1199 (1993)) as modified (A. van Hoof, P. J. Green, Plant Journal **10**:415-424 (1996)).

Initial screening gave rise to two transgenic lines, A6 and B16, that accumulated *CBF1* transcripts at elevated levels.

Figure 8 is a Northern blot showing *CBF1* and *COR* transcript levels in RLD and transgenic *Arabidopsis* plants. Leaves from nonacclimated and three-day cold-acclimated plants (*Arabidopsis thaliana* ecotype RLD plants were grown in pots under continuous light (~100 μE m^{-2}•S^{-1}) at 22°C for 18-25 days as described (Gilmour, S. J., et al., Plant Physiol. **87**:735 (1988)). In some cases, plants were then cold-acclimated by placing them at 2.5°C under continuous light (~50 μE m^{-2}•S^{-1}) for varying amounts of time. Leaves were harvested and total RNA prepared and analyzed for *CBF1* and *COR* transcripts by RNA blot analysis using ^{32}P-radiolabeled probes (Total RNA was isolated from plant leaves and subjected to RNA blot analysis using high stringency hybridization and wash conditions as described (E.J. Stockinger, et al., Proc. Natl. Acad. Sci. USA **94**:1035 (1997); and S.J. Gilmour, et al., Plant Physiol. **87**:735 (1988)).

Figure 9 is an immunoblot showing *COR15*am protein levels in RLD and transgenic *Arabidopsis* plants. Total soluble protein (100 μg) was prepared from leaves of the nonacclimated RLD (RLDw), 4-day cold-acclimated RLD (RLDc4), 7-day cold-acclimated RLD (RLDc7) and nonacclimated A6 and B16 plants and the levels of *COR15*am determined by immunoblot analysis using antiserum raised against the *COR15*am polypeptide (Total soluble protein was isolated from plant leaves, fractionated by tricine SDS-PAGE and transferred to 0.2 micron nitrocellulose as previously described (N. N. Artus et al., Proc. Natl. Acad. Sci. U.S.A. **93**:13404 (1996)). COR15am protein was detected using antiserum raised to purified COR15am and protein A conjugated alkaline phosphatase (Sigma, St. Louis, MO) (N. N. Artus et al., Proc. Natl. Acad. Sci. U.S.A. **93**:13404 (1996)). No reacting bands were observed with preimmune serum (not shown).

Southern analysis indicated that the A6 line had a single DNA insert while the B16 line had multiple inserts (not shown). Examination of fourth generation homozygous A6 and B16 plants indicated that *CBF1* transcript levels were higher in nonacclimated A6 and B16 plants than they were in nonacclimated RLD plants, the levels in A6 being about three fold higher than in B16 (Figure 8).

CBF1 overexpression resulted in strong induction of *COR* gene expression (Figure 8). Specifically, the transcript levels of *COR6.6, COR15a, COR47* and *COR78* were dramatically elevated in nonacclimated A6 and B16 plants as compared to nonacclimated RLD plants. The effect was greater in the A6 line, where *COR* transcript levels in nonacclimated plants approximated those found in cold-acclimated RLD plants. The finding that *COR* gene expression was greater in A6 plants than in B16 plants was consistent with *CBF1* transcript levels being higher in the A6 plants (Figure 7A). Immunoblot analysis indicated that the levels of the COR15am (Figure 9) and COR6.6 (not shown) polypeptides were also elevated in the A6 and B16 lines, the level of expression again being higher in the A6 line. Attempts to identify the CBF1 protein in either RLD or transgenic plants were unsuccessful. Overexpression of *CBF1* had no effect on the transcript levels for *eIF4A* (eukaryotic initiation factor 4A) (Metz, A.M., et al., Gene **120**:313 (1992)), a constitutively expressed gene that is not responsive to low temperature (Figure 8) and had no obvious effects on plant growth and development.

The results from this example demonstrate that overexpression of the *Arabidopsis* transcriptional activator CBF1 induces expression of an *Arabidopsis COR* "regulon" composed of genes carrying the CRT/DRE DNA regulatory element. It appears that CBF1 binds to the CRT/DRE DNA regulatory elements present in the promoters of these genes and activates transcription which is consistent with the notion of CBF1 having a role in *COR* gene regulation. Significantly, there was a strong correlation between CBF1 transcript levels and the magnitude of *COR* gene induction in nonacclimated A6, B16, and RLD plants (Figure 8). However, upon low temperature treatment the level of CBF1 transcripts remained relatively low in RLD plants, while *COR* gene expression was induced to about the same level as that in nonacclimated A6 plants (Figure 8). Thus, it appears that CBF1 or an associated protein becomes "activated" in response to low temperature.

3. *CBF1* overexpression resulted in a marked increase in plant freezing tolerance

The following example describes a comparison of the freezing tolerance of nonacclimated *Arabidopsis* plants which overexpress CBF1 to that of cold-

acclimated wild-type plants. As described below, the freezing tolerance of nonacclimated *Arabidopsis* plants overexpressing CBF1 significantly exceeded that of non-acclimated wild-type *Arabidopsis* plants and approached that of cold-acclimated wild-type plants.

Freezing tolerance was determined using the electrolyte leakage test (Sukumaran, N. P., et al., HortScience **7**:467 (1972)). Detached leaves were frozen to various subzero temperatures and, after thawing, cellular damage (due to freeze-induced membrane lesions) was estimated by measuring ion leakage from the tissues.

Figures 10A and 10B are graphs showing freezing tolerance of leaves from RLD and transgenic *Arabidopsis* plants. Leaves from nonacclimated RLD (RLDw) plants, cold-acclimated RLD (RLDc) plants and nonacclimated A6, B16 and T8 plants were frozen at the indicated temperatures and the extent of cellular damage was estimated by measuring electrolyte leakage (Electrolyte leakage tests were conducted as described (N.P.Sukumaran, et al., HortScience 7, 467 (1972); and S.J. Gilmour, et al., Plant Physiol. **87**:735 (1988)) with the following modifications. Detached leaves (2-4) from nonacclimated or cold-acclimated plants were placed in a test tube and submerged for 1 hour in a -2°C water-ethylene glycol bath in a completely randomized design, after which ice crystals were added to nucleate freezing. After an additional hour of incubation at -2°C, the samples were cooled in decrements of 1°C each hour until -8°C was reached. Samples (five replicates for each data point) were thawed overnight on ice and incubated in 3 ml distilled water with shaking at room temperature for 3 hours. Electrolyte leakage from leaves was measured with a conductivity meter. The solution was then removed, the leaves frozen at -80°C (for at least one hour), and the solution returned to each tube and incubated for 3 hours to obtain a value for 100% electrolyte leakage. In Figures 10A and 10B, the RLDc plants were cold-acclimated for 10 and 11 days, respectively. Error bars indicate standard deviations.

As can be seen from Figure 10A and 10B, *CBF1* overexpression resulted in a marked increase in plant freezing tolerance. The experiment presented in Figure 10A indicates that the leaves from both nonacclimated A6 and B16 plants were more freezing tolerant than those from nonacclimated RLD plants. Indeed, the

freezing tolerance of leaves from nonacclimated A6 plants approached that of leaves from cold-acclimated RLD plants. The results also indicate that the leaves from nonacclimated A6 plants were more freezing tolerant than those from nonacclimated B16 plants, a result that is consistent with the greater level of *CBF1* and *COR* gene expression in the A6 line.

The results presented in Figure 10B further demonstrate that the freezing tolerance of leaves from nonacclimated A6 plants was greater than that of leaves from nonacclimated RLD plants and that it approached the freezing tolerance of leaves from cold-acclimated RLD plants. In addition, the results indicate that overexpression of *CBF1* increases freezing tolerance to a much greater extent than overexpressing *COR15a* alone. This conclusion comes from comparing the freezing tolerance of leaves from nonacclimated A6 and T8 plants (Figure 10B). T8 plants (Artus, N. N., et al., Proc. Natl. Acad. Sci. U.S.A. **93**:13404 (1996)) are from a transgenic line that constitutively expresses *COR15a* (under control of the CaMV 35*S* promoter) at about the same level as in A6 plants (Figure 1). However, unlike in A6 plants, other CRT/DRE-regulated *COR* genes are not constitutively expressed in T8 plants (Figure 8).

A comparison of EL_{50} values (the freezing temperature that results in release of 50% of tissue electrolytes) of leaves from RLD, A6, B16 and T8 plants is presented in Table 2.

EL_{50} values were calculated and compared by analysis of variance (Model curves fitting up to third order linear polynomial trends were determined for each electrolyte leakage experiment. To insure unbiased predictions of electrolyte leakage, trends significantly improving the model fit at the 0.2 probability level were retained. EL_{50} values were calculated from the fitted models. In Table 2, an unbalanced one-way analysis of variance, adjusted for the different numbers of EL_{50} values for each plant type, was determined using SAS PROC GLM [SAS Institute, Inc. (1989), SAS/STAT User's Guide, Version 6, Cory, NC)]). EL_{50} values ± SE (n) are presented on the diagonal line for leaves from nonacclimated RLD (RLDw), cold-acclimated (7 to 10 days) RLD (RLDc) and nonacclimated A6, B16 and T8 plants. P values for comparisons of EL_{50} values are indicated in the intersecting cells.

TABLE 2

EL_{50} values					
	RLDw	**RLDc**	**A6**	**B16**	**T8**
RLDw	-3.9±0.21 (8)	P<0.0001	P<0.0001	P=0.0014	P=0.7406
RLDc		-7.6 ±0.30 (4)	P=0.3261	P<0.0001	P<0.0001
A6			-7.2±0.25 (6)	P<0.0001	P<0.0001
B16				-5.2±0.27 (5)	P=0.0044
T8					-3.8±0.35 (3)

The data confirm that: 1) the freezing tolerance of leaves from both nonacclimated A6 and B16 plants is greater than that of leaves from both nonacclimated RLD and T8 plants; and 2) that leaves from nonacclimated A6 plants are more freezing tolerant than leaves from nonacclimated B16 plants. No significant difference was detected in EL_{50} values for leaves from nonacclimated A6 and cold-acclimated RLD plants or from nonacclimated RLD and T8 plants.

The enhancement of freezing tolerance in the A6 line was also apparent at the whole plant level. Figure 11 is a photograph showing freezing survival of RLD and A6 *Arabidopsis* plants. Nonacclimated (WARM) RLD and A6 plants and 5-day cold-acclimated (COLD) RLD plants were frozen at -5°C for 2 days and then returned to a growth chamber at 22°C (Pots (3.5 inch) containing about 40 nonacclimated *Arabidopsis* plants (20 day old) and 4 day cold-acclimated plants (25 days old) (*Arabidopsis thaliana* ecotype RLD plants were grown in pots under continuous light (~100 μE m^{-2}•s^{-1}) at 22°C for 18-25 days as described (S.J. Gilmour, et al., Plant Physiol. **87**:735 (1988)). In some cases, plants were then cold-acclimated by placing them at 2.5°C under continuous light (~50 μE m^{-2}•s^{-1}) for varying amounts of time) were placed in a completely randomized design in a -5°C cold chamber in the dark. After 1 hour, ice chips were added to each pot to nucleate

freezing. Plants were removed after 2 days and returned to a growth chamber at 22°C.). A photograph of the plants after 7 days of regrowth is shown.

Although the magnitude of the difference varied from experiment to experiment, nonacclimated A6 plants consistently displayed greater freezing tolerance in whole plant freeze tests than did nonacclimated RLD plants (Figure 11). No difference in whole plant freeze survival was detected between nonacclimated B16 and RLD plants or nonacclimated T8 and RLD plants (not shown).

The results of this experiment show that CBF1-induced expression of CRT/DRE-regulated *COR* genes result in a dramatic increase in freezing tolerance and confirms the belief that COR genes play a major role in plant cold acclimation. The increase in freezing tolerance brought about by expressing the battery of CRT/DRE-regulated *COR* genes was much greater than that brought about by overexpressing *COR15a* alone indicating that *COR* genes in addition to *COR15a* have roles in freezing tolerance.

Traditional plant breeding approaches have met with limited success in improving the freezing tolerance of agronomic plants (Thomashow, M. F., Adv. Genet **28**:99 (1990)). For instance, the freezing tolerance of the best wheat varieties today is essentially the same as the most freezing-tolerance varieties developed in the early part of this century. Thus, in recent years there has been considerable interest that biotechnology might offer new strategies to improve the freezing tolerance of agronomic plants. By the results of the present invention, Applicants demonstrate the ability to enhance the freezing tolerance of nonacclimated *Arabidopsis* plants by increasing the expressing of the *Arabidopsis* regulatory gene *CBF1*. As described throughout this application, the ability of the present invention to modify the expression of environmental stress tolerance genes such as core genes has wide ranging implications since the CRT/DRE DNA regulatory element is not limited to *Arabidopsis* (Jiang C., et al., Plant Mol. Biol. **30**:679 (1996)). Rather, CBF1 and homologous genes can be used to manipulate expression of CRT/DRE-regulated *COR* genes in important crop species and thereby improve their freezing tolerance. By transforming modified versions of *CBF1* (or homologs) into such plants, it will extend their safe growing season, increase yield and expand areas of production.

3. Identification of CBF1 Homologs CBF2 and CBF3 Using CBF1

This example describes two homologs of CBF1 from *Arabidopsis thaliana* and named them CBF2 and CBF3.

CBF2 and *CBF3* have been cloned and sequenced as described below. The sequences of the DNA and encoded proteins are set forth in SEQ ID NOS: 12, 13, 14 and 15. Figure 12 shows the DNA sequence for *CBF2* encoding CBF2. Figure 13 shows the DNA sequence for *CBF3* encoding CBF3.

A lambda cDNA library prepared from RNA isolated from *Arabidopsis thaliana* ecotype Columbia (Lin and Thomashow, Plant Physiol. **99**: 519-525 (1992)) was screened for recombinant clones that carried inserts related to the *CBF1* gene (Stockinger, E. J., et al., Proc Natl Acad Sci USA **94**:1035-1040 (1997)). CBF1 was ^{32}P-radiolabeled by random priming (Sambrook et al., Molecular Cloning. A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, New York (1989)) and used to screen the library by the plaque-lift technique using standard stringent hybridization and wash conditions (Hajela, R. K., et al., Plant Physiol **93**:1246-1252 (1990); Sambrook et al., Molecular Cloning. A Laboratory Manual, Ed 2. Cold Spring Harbor laboratory Press, New York (1989) 6 X SSPE buffer, 60°C for hybridization and 0.1 X SSPE buffer and 60°C for washes). Twelve positively hybridizing clones were obtained and the DNA sequences of the cDNA inserts were determined at the MSU-DOE Plant Research Laboratory sequencing facility. The results indicated that the clones fell into three classes. One class carried inserts corresponding to *CBF1*. The two other classes carried sequences corresponding to two different homologs of *CBF1*, designated *CBF2* and *CBF3*. The nucleic acid sequences and predicted protein coding sequences for *CBF1*, *CBF2* and *CBF3* appear at Figure14.

A comparison of the nucleic acid sequences of *CBF1*, *CBF2* and *CBF3* indicate that they are 83 to 85% identical as shown in Table 3. Figure 14 shows the amino acid alignment of proteins CBF1, CBF2 and CBF3.

TABLE 3

	Percent identity[a]	
	DNA[b]	Polypeptide

cbf1/cbf2	85	86
cbf1/cbf3	83	84
cbf2/cbf3	84	85

[a] Percent identity was determined using the *Clustal* algorithm from the Megalign program (DNASTAR, Inc.).

[b] Comparisons of the nucleic acid sequences of the open reading frames are shown.

Similarly, the amino acid sequences of the three CBF polypeptides range from 84 to 86% identity. An alignment of the three amino acidic sequences reveals that most of the differences in amino acid sequence occur in the acidic C-terminal half of the polypeptide. This region of CBF1 serves as an activation domain in both yeast and *Arabidopsis* (not shown).

Residues 47 to 106 of CBF1 correspond to the AP2 domain of the protein, a DNA binding motif that to date, has only been found in plant proteins. A comparison of the AP2 domains of CBF1, CBF2 and CBF3 indicates that there are a few differences in amino acid sequence. These differences in amino acid sequence might have an effect on DNA binding specificity.

4. Activation of Transcription In Yeast Containing C-repeat/DRE Using CBF1, CBF2 and CBF3

This example shows that CBF1, CBF2 and CBF3 activate transcription in yeast containing CRT/DREs upstream of a reporter gene. The CBFs were expressed in yeast under control of the ADC1 promoter on a 2μ plasmid (pDB20.1; Berger, S. L., et al., Cell **70**:251-265 (1992)). Constructs expressing the different CBFs were transformed into yeast reporter strains which had the indicated CRT/DRE upstream of the lacZ reporter gene. Copy number of the CRT/DREs and its orientation relative to the direction of transcription from each promoter is indicated by the direction of the arrow.

Figure 15 is a graph showing transcription regulation of CRT/DRE containining reporter genes by *CBF1*, *CBF2* and *CBF3* genes in yeast. In Figure 15, the vertical lines across the arrows of the COR15a construct represent the m3cor15a mutant CRT/DRE construct. Each CRT/DRE-lacZ construct was integrated into the URA3 locus of yeast. Error bars represent the standard deviation derived from three replicate transformation events with the same CBF activator construct into the respective reporter strain. Quantitative B-gal assays were performed as described by Rose and Botstein (Rose, M., et al., Methods Enzymol. **101**:167-180 (1983)).

5. Homologous CBF Encoding Genes In Other Plants.

This example shows that homologous sequences to CBF1 are present in other plants. The presence of these homologous sequences suggest that the same or similar cold regulated binding domains such as the C-repeat/DRE of Arabidopsis (CCGAC) exist in other plants. This example serves to indicate that genes with significant homology to CBF1, CBF2 and CBF3 exist in a wide range of plant species.

Total plant DNAs from *Arabadopsis thaliana*, *Nicotiana tabacum*, *Lycopersicon pimpinellifolium*, *Prunus avium*, *Prunus cerasus*, *Cucumis sativus*, and *Oryza sativa* were isolated according to Stockinger al (Stockinger, E. J., et al., J. Heredity, **87**:214-218 (1996)). Approximately 2 to 10 µg of each DNA sample was restriction digested, transferred to nylon membrane (Micron Separations, Westboro, MA) and hybridized according to Walling et al. (Walling, L. L., et al., Nucleic Acids Res. **16**:10477-10492 (1988)). Hybridization conditions were: 42∘C in 50% formamide, 5X SSC, 20 mM phosphate buffer 1X Denhardt's, 10% dextran sulfate, and 100µg/ml herring sperm DNA. Four low stringency washes at RT in 2X SSC, 0.05% Na sarcosyl and 0.02% Na_4 pyrophosphate were performed prior to high stringency washes at 55∘C in 0.2X SSC, 0.05% Na sarcosyl and 0.01% Na_4 pyrophosphate. High stringency washes were performed until no counts were detected in the washout. The BclI-BglII fragment of CBF1 (Stockinger et al., Proc Natl Acad Sci USA **94**:1035-1040 (1997)) was gel isolated (Sambrook et al., Molecular Cloning. A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, New York (1989)) and direct prime labelled (Feinberg and Vogelstein, Anal. Biochem **132**: 6-13 (1982)) using the primer MT117 (TTGGCGGCTACGAATCCC; SEQ ID NO:16). Specific activity of the radiolabelled fragment was approximately 4 x 10^8 cpm/µg. Autoradiography was performed using HYPERFILM-MP (Amersham) at -80∘C with one intensifying screen for 15 hours.

Autoradiography of the gel showed that DNA sequences from *Arabadopsis thaliana*, *Nicotiana tabacum*, *Lycopersicon pimpinellifolium*, *Prunis avium*, *Prunus cerasus*, *Cucumis sativus*, and *Oryza sativa* hybridized to the labeled BclI, BglII fragment of *CBF1*. These results suggest that homologous CBF encoding genes are present in a variety of other plants.

6. Identification Of Homologous Sequence To CBF1 In Canola

This example describes the identification of homologous sequences to CBF1 in canola using PCR. Degenerate primers were designed for regions of AP2 binding domain and outside of the AP2 (carboxyl terminal domain). More specifically, the following degenerate PCR primers were used:

Mol 368 (reverse) 5'- CAY CCN ATH TAY MGN GGN GT -3'

Mol 378 (forward) 5'- GGN ARN ARC ATN CCY TCN GCC -3'

(Y: C/T, N: A/C/G/T, H: A/C/T, M: A/C, R: A/G)

Primer Mol 368 is in the AP2 binding domain of CBF1 (amino acid seq: H P I Y R G V) while primer Mol 378 is outside the AP2 domain (carboxyl terminal domain)(amino acid seq: M A E G M L L P).

PCR was genomic DNA isolated from *Brassica Napus* using these primers as follows: an initial denaturation step of 2 min at 93 °C; 35 cycles of 93 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min ; and a final incubation of 7 min at 72 °C at the end of cycling.

The PCR products were separated by electrophoresis on a 1.2% agarose gel and, transferred to nylon membrane and hybridized with the AT CBF1 probe prepared from Arabidopsis genomic DNA by PCR amplification. The hybridized products were visualized by colormetric detection system (Boehringer Mannheim) and the corresponding bands from a similar agarose gel were isolated (By Qiagen Extraction Kit). The DNA fragments were ligated into the TA clone vector from TOPO TA Cloning Kit (Invitrogen) and transformed into E. coli strain TOP10 (Invitrogen).

7 colonies were picked and the inserts were sequenced on an ABI 377 machine from both strands of sense and antisense after plasmid DNA isolation. The DNA sequence was edited by sequencer and aligned with the AtCBF1 by GCG software and NCBI blast searching.

Figure 16 shows an amino acid sequence of a homolog [CAN1; SEQ. ID. No. 17] identified by this process and its alignment to the amino acid sequence of CBF1. The nucleic acid sequence for CAN1 is listed herein as SEQ. ID. No. 18.

As illustrated in Figure 16, the DNA sequence alignment in four regions of BN-CBF1 shows 82% identity in the AP2 binding domain region and range from 75% to 83% with some alignment gaps due to regions of lesser homology or introns in the genomic sequence. The aligned amino acid sequences show that the BNCBF1 gene has 88% identity in the AP2 domain region and 85% identity outside the AP2 domain when aligned for two insertion sequences that are outside the AP2 domain. The extra amino acids in the 2 insertion regions are either due to the presence of introns in this region of the BNCBF1 gene, as it was derived from genomic DNA, or could be due to extra amino acids in these regions of the BNCBF1 gene. Isolation and sequencing of a cDNA of the BNCBF1 gene using the genomic DNA as a probe will resolve this.

While the present invention is disclosed by reference to the preferred embodiments and examples detailed above, it is to be understood that these examples are intended in an illustrative rather than limiting sense, as it is contemplated that modifications will readily occur to those skilled in the art, which modifications will be within the spirit of the invention and the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Michael F. Thomashow, Eric J. Stockinger and Cai-Zhong Jiang

(ii) TITLE OF INVENTION: TRANSFORMED PLANT WITH MODIFIED ENVIRONMENTAL STRESS TOLERANCE GENE EXPRESSION

(iii) NUMBER OF SEQUENCES: 18

(A) ADDRESSEE: David J. Weitz, Wilson Sonsini Goodrich & Rosati
(B) STREET: 650 Page Mill Road
(C) CITY: Palo Alto
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94304-1050

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: 3.5 inch diskette
(B) COMPUTER: IBM compatible
(C) OPERATING SYSTEM: Microsoft Windows 95
(D) SOFTWARE: Wordperfect for windows 6.0, ASCII (DOS) TEXT format

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/706,270
(B) FILING DATE: September 4, 1996

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: David J. Weitz
(B) REGISTRATION NUMBER: 38,362
(C) REFERENCE/DOCKET NUMBER: 19117.706

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (650) 493-9300
(B) TELEFAX: (650) 493-6811
(C) TELEX: None

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 905
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Arabidopsis thaliana*
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY: CBF1 gene
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
AAAAAGAATC TACCTGAAAA GAAAAAAAAG AGAGAGAGAT ATAAATAGCT   50

TACCAAGACA GATATACTAT CTTTTATTAA TCCAAAAAGA CTGAGAACTC  100

TAGTAACTAC GTACTACTTA AACCTTATCC AGTTTCTTGA AACAGAGTAC  150

TCTGATCAAT GAACTCATTT TCAGCTTTTT CTGAAATGTT TGGCTCCGAT  200

TACGAGCCTC AAGGCGGAGA TTATTGTCCG ACGTTGGCCA CGAGTTGTCC  250

GAAGAAACCG GCGGGCCGTA AGAAGTTTCG TGAGACTCGT CACCCAATTT  300

ACAGAGGAGT TCGTCAAAGA AACTCCGGTA AGTGGGTTTC TGAAGTGAGA  350

GAGCCAAACA AGAAAACCAG GATTTGGCTC GGGACTTTCC AAACCGCTGA  400

GATGGCAGCT CGTGCTCACG ACGTCGCTGC ATTAGCCCTC CGTGGCCGAT  450

CAGCATGTCT CAACTTCGCT GACTCGGCTT GGCGGCTACG AATCCCGGAG  500

TCAACATGCG CCAAGGATAT CCAAAAAGCG GCTGCTGAAG CGGCGTTGGC  550

TTTTCAAGAT GAGACGTGTG ATACGACGAC CACGGATCAT GGCCTGGACA  600

TGGAGGAGAC GATGGTGGAA GCTATTTATA CACCGGAACA GAGCGAAGGT  650
```

GCGTTTTATA TGGATGAGGA GACAATGTTT GGGATGCCGA CTTTGTTGGA 700

TAATATGGCT GAAGGCATGC TTTTACCGCC GCCGTCTGTT CAATGGAATC 750

ATAATTATGA CGGCGAAGGA GATGGTGACG TGTCGCTTTG GAGTTACTAA 800

TATTCGATAG TCGTTTCCAT TTTTGTACTA TAGTTTGAAA ATATTCTAGT 850

TCCTTTTTTA GAATGGTTCC TTCATTTTAT TTTATTTTAT TGTTGTAGAA 900

ACGAG 905

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 213
(B) TYPE: Amino Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Arabidopsis thaliana*
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY: CBF1 protein
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asn Ser Phe Ser Ala Phe Ser Glu Met Phe Gly Ser Asp Tyr

```
                    5                   10                  15
Glu Pro Gln Gly Gly Asp Tyr Cys Pro Thr Leu Ala Thr Ser Cys
                   20                  25                  30
Pro Lys Lys Pro Ala Gly Arg Lys Lys Phe Arg Glu Thr Arg His
                   35                  40                  45
Pro Ile Tyr Arg Gly Val Arg Gln Arg Asn Ser Gly Lys Trp Val
                   50                  55                  60
Ser Glu Val Arg Glu Pro Asn Lys Lys Thr Arg Ile Trp Leu Gly
                   65                  70                  75
Thr Phe Gln Thr Ala Glu Met Ala Ala Arg Ala His Asp Val Ala
                   80                  85                  90
Ala Leu Ala Leu Arg Gly Arg Ser Ala Cys Leu Asn Phe Ala Asp
                   95                 100                 105
Ser Ala Trp Arg Leu Arg Ile Pro Glu Ser Thr Cys Ala Lys Asp
                  110                 115                 120
Ile Gln Lys Ala Ala Ala Glu Ala Ala Leu Ala Phe Gln Asp Glu
                  125                 130                 135
Thr Cys Asp Thr Thr Thr Thr Asp His Gly Leu Asp Met Glu Glu
                  140                 145                 150
Thr Met Val Glu Ala Ile Tyr Thr Pro Glu Gln Ser Glu Gly Ala
                  155                 160                 165
Phe Tyr Met Asp Glu Glu Thr Met Phe Gly Met Pro Thr Leu Leu
                  170                 175                 180
Asp Asn Met Ala Glu Gly Met Leu Leu Pro Pro Pro Ser Val Gln
                  185                 190                 195
Trp Asn His Asn Tyr Asp Gly Glu Gly Asp Gly Asp Val Ser Leu
                  200                 205                 210
Trp Ser Tyr
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: N/A - Synthetic
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A

(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GATCATTTCA TGGCCGACCT GCTTTTT 27

(3) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: N/A - Synthetic
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CACAATTTCA AGAATTCACT GCTTTTTT 28

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: N/A - Synthetic
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GATCATTTCA TGGTATGTCT GCTTTTT 27

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: N/A - Synthetic
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GATCATTTCA TGGAATCACT GCTTTTT 27

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: N/A - Synthetic
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GATCACTTGA TGGCCGACCT CTTTTTT 27

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: N/A - Synthetic
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GATCAATATA CTACCGACAT GAGTTCT 27

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: N/A - Synthetic
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ACTACCGACA TGAGTTCCAA AAAGC 25

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60
(B) TYPE: Amino Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Amino Acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Arabidopsis thaliana*
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION: Figure 2D

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Ile Tyr Arg Gly Val Arg Gln Arg Asn Ser Gly Lys Trp Val Ser
                 5                  10                  15
Glu Val Arg Glu Pro Asn Lys Lys Thr Arg Ile Trp Leu Gly Thr
                20                  25                  30
Phe Gln Thr Ala Glu Met Ala Ala Arg Ala His Asp Val Ala Ala
                35                  40                  45
Leu Ala Leu Arg Gly Arg Ser Ala Cys Leu Asn Phe Ala Asp Ser
                50                  55                  60
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61
(B) TYPE: Amino Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Polypeptide

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Tobacco
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION: Figure 2D

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
His Tyr Arg Gly Val Arg Gln Arg Pro Trp Gly Lys Phe Ala Ala
                 5                  10                  15
Glu Ile Arg Asp Pro Ala Lys Asn Gly Ala Arg Val Trp Leu Gly
            20                  25                  30
Thr Tyr Glu Thr Ala Glu Glu Ala Ala Leu Ala Tyr Asp Lys Ala
            35                  40                  45
Ala Tyr Arg Met Arg Gly Ser Lys Ala Leu Leu Asn Phe Pro His
            50                  55                  60
Arg
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 651
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Arabidopsis thaliana*
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY: CBF2
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
ATGAACTCAT TTTCTGCCTT TTCTGAAATG TTTGGCTCCG ATTACGAGTC TCCGGTTTCC  60
TCAGGCGGTG ATTACAGTCC GAAGCTTGCC ACGAGCTGCC CCAAGAAACC AGCGGGAAGG 120
AAGAAGTTTC GTGAGACTCG TCACCCAATT TACAGAGGAG TTCGTCAAAG AAACTCCGGT 180
AAGTGGGTGT GTGAGTTGAG AGAGCCAAAC AAGAAAACGA GGATTTGGCT CGGGACTTTC 240
CAAACCGCTG AGATGGCAGC TCGTGCTCAC GACGTCGCCG CCATAGCTCT CCGTGGCAGA 300
TCTGCCTGTC TCAATTTCGC TGACTCGGCT TGGCGGCTAC GAATCCCGGA ATCAACCTGT 360
GCCAAGGAAA TCCAAAAGGC GGCGGCTGAA GCCGCGTTGA ATTTTCAAGA TGAGATGTGT 420
CATATGACGA CGGATGCTCA TGGTCTTGAC ATGGAGGAGA CCTTGGTGGA GGCTATTTAT 480
ACGCCGGAAC AGAGCCAAGA TGCGTTTTAT ATGGATGAAG AGGCGATGTT GGGGATGTCT 540
AGTTTGTTGG ATAACATGGC CGAAGGGATG CTTTTACCGT CGCCGTCGGT TCAATGGAAC 600
TATAATTTTG ATGTCGAGGG AGATGATGAC GTGTCCTTAT GGAGCTATTA A          651
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 216

(B) TYPE: Amino Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Polypeptide

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Arabidopsis thaliana*
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY: CBF2
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Met Asn Ser Phe Ser Ala Phe Ser Glu Met Phe Gly Ser Asp Tyr
                5                   10                  15
Glu Ser Pro Val Ser Ser Gly Gly Asp Pyr Ser Pro Lys Leu Als
                20                  25                  30
Thr Ser Cys Pro Lys Lys Pro Ala Gly Arg Lys Lys Phe Arg Glu
                35                  40                  45
Thr Arg His Pro Ile Tyr Arg Gly Val Arg Gln Arg Asn Ser Gly
                50                  55                  60
Lys Trp Val Cys Glu Leu Arg Glu Pro Asn Lys Lys Thr Arg Ile
                65                  70                  75
Trp Leu Gly Thr Phe Gln Thr Ala Glu Met Ala Ala Arg Ala His
                80                  85                  90
Asp Val Ala Ala Ile Ala Leu Arg Gly Arg Ser Ala Cys Leu Asn
                95                  100                 105
Phe Ala Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Ser Thr Cys
                110                 115                 120
Ala Lys Glu Ile Gln Lys Ala Ala Ala Glu Ala Ala Leu Asn Phe
                125                 130                 135
```

```
Gln Asp Glu Met Cys His Met Thr Thr Asp Ala His Gly Leu Asp
                140                 145                 150
Met Glu Glu Thr Leu Val Glu Ala Ile Tyr Thr Pro Glu Gln Ser
                155                 160                 165
Gln Asp Ala Phe Tyr Met Asp Glu Glu Ala Met Leu Gly met Ser
                170                 175                 180
Ser Leu Leu Asp Asn Met Ala Glu Gly Met Leu Leu Pro Ser Pro
                185                 190                 195
Ser Val Gln Trp Asn Tyr Asn Phe Asp Val Glu Gly Asp Asp Asp
                200                 205                 210
Val Ser Leu Trp Ser Tyr
                215
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 651
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Arabidopsis thaliana*
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY: CBF3
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
ATGAACTCAT TTTCTGCTTT TTCTGAAATG TTTGGCTCCG ATTACGAGTC TTCGGTTTCC 60
TCAGGCGGTG ATTATATTCC GACGCTTGCG AGCAGCTGCC CCAAGAAACC GGCGGGTCGT 120
AAGAAGTTTC GTGAGACTCG TCACCCAATA TACAGAGGAG TTCGTCGGAG AAACTCCGGT 180
AAGTGGGTTT GTGAGGTTAG AGAACCAAAC AAGAAAACAA GGATTTGGCT CGGAACATTT 240
CAAACCGCTG AGATGGCAGC TCGAGCTCAC GACGTTGCCG CTTTAGCCCT TCGTGGCCGA 300
TCAGCCTGTC TCAATTTCGC TGACTCGGCT TGGAGACTCC GAATCCCGGA ATCAACTTGC 360
GCTAAGGACA TCCAAAAGGC GGCGGCTGAA GCTGCGTTGG CGTTTCAGGA TGAGATGTGT 420
GATGCGACGA CGGATCATGG CTTCGACATG GAGGAGACGT TGGTGGAGGC TATTTACACG 480
GCGGAACAGA GCGAAAATGC GTTTTATATG CACGATGAGG CGATGTTTGA GATGCCGAGT 540
TTGTTGGCTA ATATGGCAGA AGGGATGCTT TTGCCGCTTC CGTCCGTACA GTGGAATCAT 600
AATCATGAAG TCGACGGCGA TGATGACGAC GTATCGTTAT GGAGTTATTA A          651
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 216
(B) TYPE: Amino Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Polypeptide

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Arabidopsis thaliana*
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY: CBF3
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Asn Ser Phe Ser Ala Phe Ser Glu Met Phe Gly Ser Asp Tyr
 1               5                  10                  15
```

```
Glu Ser Ser Val Ser Ser Gly Gly Asp Tyr Ile Pro Thr Leu Ala
                 20                  25                  30
Ser Ser Cys Pro Lys Lys Pro Ala Gly Arg Lys Lys Phe Arg Glu
                 35                  40                  45
Thr Arg His Pro Ile Tyr Arg Gly Val Arg Arg Arg Asn Ser Gly
                 50                  55                  60
Lys Trp Val Cys Glu Val Arg Glu Pro Asn Lys Lys Thr Arg Ile
                 65                  70                  75
Trp Leu Gly Thr Phe Gln Thr Ala Glu Met Ala Ala Arg Ala His
                 80                  85                  90
Asp Val Als Ala Leu Ala Leu Arg Gly Arg Ser Ala Cys Leu Asn
                 95                 100                 105
Phe Ala Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Ser Thr Cys
                110                 115                 120
Ala Lys Asp Ile Gln Lys Ala Ala Ala Glu Ala Ala Leu Ala Phe
                125                 130                 135
Gln Asp Glu Met Cys Asp Ala Thr Thr Asp His Gly Phe Asp Met
                140                 145                 150
Glu Glu Thr Leu Val Glu Ala Ile Tyr Thr Ala Glu Gln Ser Glu
                155                 160                 165
Asn Ala Phe Tyr Met His Asp Glu Ala Met Phe Glu Met Pro Ser
                170                 175                 180
Leu Leu Ala Asn Met Ala Glu Gly Met Leu Leu Pro Leu Pro Ser
                185                 190                 195
Val Gln Trp Asn His Asn His Glu Val Asp Gly Asp Asp Asp Asp
                200                 205                 210
Val Ser Leu Trp Ser Tyr
                215
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Tobacco
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A

(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD: sequencing
(D) OTHER INFORMATION: Figure 2D

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TTGGCGGCTA CGAATCCC 18

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 210
(B) TYPE: Amino Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Polypeptide

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Brassica Napus*
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A
(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
His Pro Ile Tyr Arg Gly Val Arg Leu Arg Lys Ser Gly Lys Trp
                  5                  10                  15
Val Cys Glu Val Arg Glu Pro Asn Lys Lys Ser Arg Ile Trp Leu
             20                  25                  30
Gly Thr Phe Lys Thr Ala Glu Met Ala Ala Arg Ala His Asp Val
         35                  40                  45
Ala Ala Leu Ala Leu Arg Gly Arg Gly Ala Cys Leu Asn Tyr Ala
     50                  55                  60
Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Thr Thr Cys His Lys
 65                  70                  75
Asp Ile Gln Lys Ala Ala Ala Glu Ala Ala Leu Ala Phe Glu Ala
             80                  85                  90
Glu Lys Ser Asp Val Thr Met Gln Asn Gly Gln Asn Met Glu Glu
         95                 100                 105
Thr Thr Ala Val Ala Ser Gln Ala Glu Val Asn Asp Thr Thr Thr
    110                 115                 120
Glu His Gly Met Asn Met Glu Glu Ala Thr Ala Val Ala Ser Gln
    125                 130                 135
Ala Glu Val Asn Asp Thr Thr Thr Asp His Gly Val Asp Met Glu
    140                 145                 150
Glu Thr Met Val Glu Ala Val Phe Thr Gly Glu Gln Ser Glu Gly
    155                 160                 165
Phe Asn Met Ala Lys Glu Ser Thr Val Glu Ala Ala Val Val Thr
    170                 175                 180
Glu Glu Pro Ser Lys Gly Ser Tyr Met Asp Glu Glu Trp Met Leu
    185                 190                 195
Glu Met Pro Thr Leu Leu Ala Asp Met Ala Glu Gly Met Leu Leu
    200                 205                 210
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 632
(B) TYPE: Nucleic Acid Sequence
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *Canola*
(B) STRAIN:
(C) INDIVIDUAL ISOLATE: N/A

(D) DEVELOPMENTAL STAGE: N/A
(E) HAPLOTYPE: N/A
(F) TISSUE TYPE: N/A
(G) CELL TYPE: N/A
(H) CELL LINE: N/A
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CACCCGATAT ACCGGGGAGT TCGTCTGAGA AAGTCAGGTA AGTGGGTGTG 50

TGAAGTGAGG GAACCAAACA AGAAATCTAG AATTTGGCTT GGAACTTTCA 100

AAACAGCTGA GATGGCAGCT CGTGCTCACG ACGTCGCTGC CCTAGCCCTC 150

CGTGGAAGAG GCGCCTGCCT CAATTATGCG GACTCGGCTT GGCGGCTCCG 200

CATCCCGGAG ACAACCTGCC ACAAGGATAT CCAGAAGGCT GCTGCTGAAG 250

CCGCATTGGC TTTTGAGGCT GAGAAAAGTG ATGTGACGAT GCAAAATGGC 300

CAGAACATGG AGGAGACGAC GGCGGTGGCT TCTCAGGCTG AAGTGAATGA 350

CACGACGACA GAACATGGCA TGAACATGGA GGAGGCAACG GCAGTGGCTT 400

CTCAGGCTGA GGTGAATGAC ACGACGACGG ATCATGGCGT AGACATGGAG 450

GAGACAATGG TGGAGGCTGT TTTTACTGGG GAACAAAGTG AAGGGTTTAA 500

CATGGCGAAG GAGTCGACGG TGGAGGCTGC TGTTGTTACG GAGGAACCGA 550

GCAAAGGATC TTACATGGAC GAGGAGTGGA TGCTCGAGAT GCCGACCTTG 600

TTGGCTGATA TGGCAGAAGG GATGCTCCTG CC 632

WE CLAIM:

1. A transformed plant comprising:

one or more environmental stress tolerance genes;

a DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes; and

a recombinant sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence.

2. A transformed plant according to claim 1 wherein the recombinant sequence is native to the plant.

3. A transformed plant according to claim 1 wherein the recombinant sequence is not native to the plant.

4. A transformed plant according to claim 1 wherein the recombinant sequence is heterologous relative to the DNA regulatory sequence.

5. A transformed plant according to claim 1 wherein the DNA regulatory sequence includes the subsequence CCG.

6. A transformed plant according to claim 1 wherein the DNA regulatory sequence includes the subsequence CCGAC.

7. A transformed plant according to claim 1 wherein the DNA regulatory sequence includes a subsequence selected from the group consisting of CCGAA, CCGAT, CCGAC, CCGAG, CCGTA, CCGTT, CCGTC, CCGTG, CCGCA, CCGCT, CCGCG, CCGCC, CCGGA, CCGGT, CCGGC, CCGGG, AACCG, ATCCG, ACCCG, AGCCG, TACCG, TTCCG, TCCCG, TGCCG, CACCG, CTCCG, CGCCG, CCCCG, GACCG, GTCCG, GCCCG, GGCCG, ACCGA, ACCGT, ACCGC, ACCGG, TCCGA, TCCGT, TCCGC, TCCGG, CCCGA, CCCGT, CCCGC, CCCGG, GCCGA, GCCGT, GCCGC, and GCCGG.

8. A transformed plant according to claim 1 wherein the plant includes a sequence encoding for an AP2 domain capable of binding to the DNA regulatory sequence.

9. A transformed plant according to claim 1 wherein the plant includes a sequence which is a homolog of an AP2 domain encoded by CBF1, CBF2, CBF3 or CAN1.

10. A transformed cell according to claim 1 wherein the plant includes a recombinant sequence which is a homolog of CBF1, CBF2, or CBF3.

11. A transformed plant according to claim 10 wherein the recombinant sequence has at least about 83 percent identity to at least one of CBF1, CBF2, and CBF3.

12. A transformed plant according to claim 1 wherein the recombinant sequence encodes an amino acid sequence which has substantially the same homology to CBF1, CBF2, and CBF3 as these amino acid sequences have with each other.

13. A transformed plant according to claim 1 wherein the recombinant sequence encodes a binding protein capable of regulating expression of one or more environmental stress tolerance genes in a plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes.

14. A transformed plant comprising:

one or more environmental stress tolerance genes;

a DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes;

a sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence; and

a recombination promoter which regulates expression of the sequence encoding the binding protein.

15. A transformed plant according to claim 14 wherein the promoter is not native to the plant.

16. A transformed plant according to claim 14 wherein the promoter is an inducible promoter.

17. A transformed plant according to claim 14 wherein the promoter is not regulated by an environmental stress.

18. A transformed plant according to claim 14 wherein the sequence encoding the binding protein is recombinant.

19. A transformed plant according to claim 18 wherein the recombinant sequence encoding the binding protein is native to the plant.

20. A transformed plant according to claim 18 wherein the recombinant sequence encoding the binding protein is not native to the plant.

21. A transformed plant according to claim 18 wherein the recombinant sequence is heterologous relative to the promoter.

22. A transformed plant according to claim 18 wherein the recombinant sequence is heterologous relative to the DNA regulatory sequence.

23. A transformed plant according to claim 14 wherein the DNA regulatory sequence includes the subsequence CCG.

24. A transformed plant according to claim 14 wherein the DNA regulatory sequence includes the subsequence CCGAC.

25. A transformed plant according to claim 14 wherein the DNA regulatory sequence includes a subsequence selected from the group consisting of CCGAA, CCGAT, CCGAC, CCGAG, CCGTA, CCGTT, CCGTC, CCGTG, CCGCA, CCGCT, CCGCG, CCGCC, CCGGA, CCGGT, CCGGC, CCGGG, AACCG, ATCCG, ACCCG, AGCCG, TACCG, TTCCG, TCCCG, TGCCG, CACCG, CTCCG, CGCCG, CCCCG, GACCG, GTCCG, GCCCG, GGCCG, ACCGA, ACCGT, ACCGC, ACCGG, TCCGA, TCCGT, TCCGC, TCCGG, CCCGA, CCCGT, CCCGC, CCCGG, GCCGA, GCCGT, GCCGC, and GCCGG.

26. A transformed plant according to claim 14 wherein the sequence encoding the binding protein includes a sequence encoding for an AP2 domain capable of binding to the DNA regulatory sequence.

27. A transformed plant according to claim 14 wherein the sequence encoding the binding protein includes a sequence encoding an AP2 domain which is a homolog of an AP2 domain encoded by CBF1, CBF2, CBF3 or CAN1.

28. A transformed cell according to claim 14 wherein the sequence encoding the binding protein is a homolog of CBF1, CBF2, or CBF3.

29. A transformed plant according to claim 28 wherein the sequence has at least about 83 percent identity to at least one of CBF1, CBF2, and CBF3.

30. A transformed plant according to claim 14 wherein the sequence encoding the binding protein has at least about 83 percent identity to at least one of CBF1, CBF2, and CBF3.

31. A transformed plant according to claim 14 wherein the sequence encoding the binding protein encodes an amino acid sequence which has substantially the same homology to CBF1, CBF2, and CBF3 as these amino acid sequences have with each other.

32. A transformed plant according to claim 14 wherein the sequence encodes a binding protein capable of regulating expression of one or more environmental stress tolerance genes in a plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes.

33. A transformed plant comprising:

one or more environmental stress tolerance genes;

a recombinant DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes; and

a sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence.

34. A transformed plant according to claim 33 wherein the recombinant DNA regulatory sequence is native to the plant.

35. A transformed plant according to claim 33 wherein the recombinant DNA regulatory sequence is not native to the plant.

36. A transformed plant according to claim 33 wherein the sequence encoding the binding protein is heterologous relative to the recombinant DNA regulatory sequence.

37. A transformed plant according to claim 33 wherein the sequence encoding the binding protein is recombinant.

38. A transformed plant according to claim 37 wherein the sequence encoding the binding protein is native to the plant.

39. A transformed plant according to claim 37 wherein the sequence encoding the binding protein is not native to the plant.

40. A transformed plant according to claim 33 wherein the recombinant DNA regulatory sequence includes the subsequence CCG.

41. A transformed plant according to claim 33 wherein the recombinant DNA regulatory sequence includes the subsequence CCGAC.

42. A transformed plant according to claim 33 wherein the recombinant DNA regulatory sequence includes a subsequence selected from the group consisting of CCGAA, CCGAT, CCGAC, CCGAG, CCGTA, CCGTT, CCGTC, CCGTG, CCGCA, CCGCT, CCGCG, CCGCC, CCGGA, CCGGT, CCGGC, CCGGG, AACCG, ATCCG, ACCCG, AGCCG, TACCG, TTCCG, TCCCG, TGCCG, CACCG, CTCCG, CGCCG, CCCCG, GACCG, GTCCG, GCCCG, GGCCG, ACCGA, ACCGT, ACCGC, ACCGG, TCCGA, TCCGT, TCCGC, TCCGG, CCCGA, CCCGT, CCCGC, CCCGG, GCCGA, GCCGT, GCCGC, and GCCGG.

43. A transformed plant comprising:

at least one recombinant environmental stress tolerance gene;

a DNA regulatory sequence which regulates expression of the at least one environmental stress tolerance gene; and

a sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence.

44. A transformed plant according to claim 43 wherein the recombinant environmental stress tolerance gene is native to the plant.

45. A transformed plant according to claim 43 wherein the recombinant environmental stress tolerance gene is not native to the plant.

46. A transformed plant comprising:

at least one recombinant environmental stress tolerance gene;

a DNA regulatory sequence which regulates expression of the at least one environmental stress tolerance gene; and

a recombinant binding protein expressed by the plant which is capable of selectively binding to the DNA regulatory sequence.

47. A transformed plant according to claim 46 wherein the recombinant binding protein is native to the plant.

48. A transformed plant according to claim 46 wherein the recombinant binding protein is not native to the plant.

49. A transformed plant according to claim 46 wherein the recombinant binding protein is heterologous relative to the DNA regulatory sequence.

50. A transformed plant according to claim 46 wherein the DNA regulatory sequence includes the subsequence CCG.

51. A transformed plant according to claim 46 wherein the DNA regulatory sequence includes the subsequence CCGAC.

52. A transformed plant according to claim 46 wherein the DNA regulatory sequence includes a subsequence selected from the group consisting of CCGAA, CCGAT, CCGAC, CCGAG, CCGTA, CCGTT, CCGTC, CCGTG, CCGCA, CCGCT, CCGCG, CCGCC, CCGGA, CCGGT, CCGGC, CCGGG, AACCG, ATCCG, ACCCG, AGCCG, TACCG, TTCCG, TCCCG, TGCCG, CACCG, CTCCG, CGCCG, CCCCG, GACCG, GTCCG, GCCCG, GGCCG, ACCGA, ACCGT, ACCGC, ACCGG, TCCGA, TCCGT, TCCGC, TCCGG, CCCGA, CCCGT, CCCGC, CCCGG, GCCGA, GCCGT, GCCGC, and GCCGG.

53. A transformed plant according to claim 46 wherein the recombinant binding protein includes an AP2 domain capable of binding to the DNA regulatory sequence.

54. A transformed plant according to claim 46 wherein the recombinant binding protein includes an AP2 domain which is a homolog of an AP2 domain of CBF1, CBF2, CBF3 or CAN1.

55. A transformed plant according to claim 46 wherein the recombinant binding protein includes an amino acid sequence which has substantially the same homology to CBF1, CBF2, and CBF3 as these amino acid sequences have with each other.

56. A transformed plant according to claim 46 wherein the recombinant binding protein regulates expression of one or more environmental stress tolerance genes in a plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes.

TRANSFORMED PLANT WITH MODIFIED ENVIRONMENTAL STRESS TOLERANCE GENE EXPRESSION

Inventors: Michael F. Thomashow, Eric J. Stockinger and Cai-Zhong Jiang

ABSTRACT

A transformed plant with modified environmental stress tolerance gene expression is provided. The transformed plant includes one or more environmental stress tolerance genes; a DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes; and a sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence where at least one of the environmental stress tolerance genes, DNA regulatory sequence and sequence encoding a binding protein is recombinant. The DNA regulatory sequence may optionally be a member of a class of DNA regulatory sequences which includes the subsequence CCG or more particularly one of the following subsequences: CCGAA, CCGAT, CCGAC, CCGAG, CCGTA, CCGTT, CCGTC, CCGTG, CCGCA, CCGCT, CCGCG, CCGCC, CCGGA, CCGGT, CCGGC, CCGGG, AACCG, ATCCG, ACCCG, AGCCG, TACCG, TTCCG, TCCCG, TGCCG, CACCG CCCG, GACCG, GTCCG, GCCCG, GGCCG, ACCGA, ACCGT, ACCGC, ACCGG, TCCGA, TCCGT, TCCGC, TCCGG, CCCGA, CCCGT, CCCGC, CCCGG, GCCGA, GCCGT, GCCGC, and GCCGG.

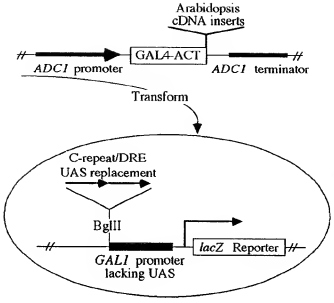

FIGURE 1A

Activity of "positive" plasmids in reporter strains

UAS Replacement Sequence			Yeast colony color on X-gal filters
Oligo	C-repeat/DRE	Inserts	
MT50	*COR15a*	►►►►►►◄◄	Blue
MT50	*COR15a*	►►◄◄◄◄◄◄	Blue
MT66	*COR78*	◄◄►►►	Blue
MT52	M1 *COR15a*	►◄◄◄	White

FIGURE 1B

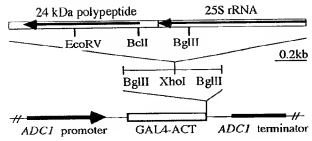

FIGURE 2A

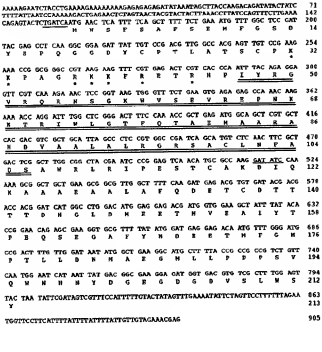

FIGURE 2B

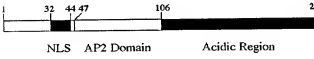

FIGURE 2C

FIGURE 2D

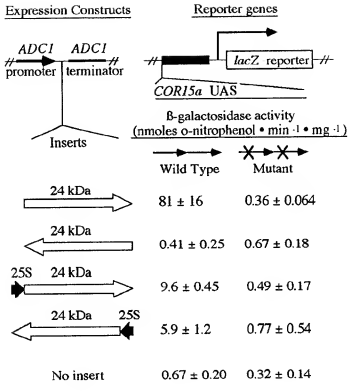
Expression Constructs
Reporter genes
ADC1
promoter
ADC1
terminator
lacZ reporter
COR15a UAS
Inserts
ß-galactosidase activity
(nmoles o-nitrophenol • min -1 • mg -1)
Wild Type
Mutant
24 kDa
81 ± 16
0.36 ± 0.064
24 kDa
0.41 ± 0.25
0.67 ± 0.18
25S
24 kDa
9.6 ± 0.45
0.49 ± 0.17
24 kDa
25S
5.9 ± 1.2
0.77 ± 0.54
No insert
0.67 ± 0.20
0.32 ± 0.14

FIGURE 3

FIGURE 4

BSA
Sense insert
Antisense insert
Vector
Competitor DNAs
COR15a 1X
COR15a 2X
M1COR15a 2X
M2COR15a 2X
M3COR15a 2X
COR78-1 2X
COR78-1 3X
COR78-2 2X

FIGURE 5

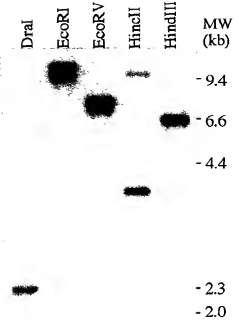

FIGURE 6

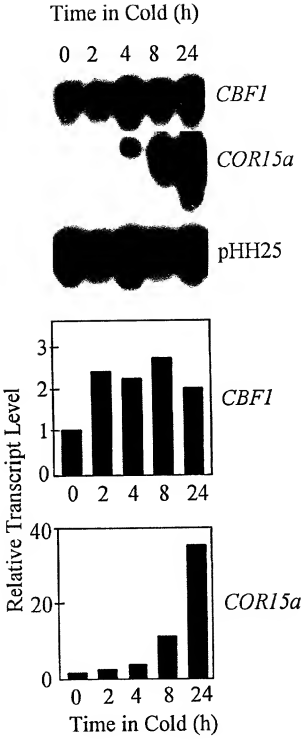

FIGURE 7

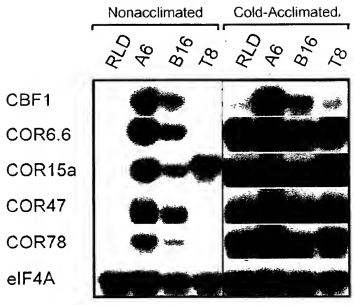

FIGURE 8

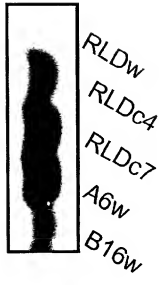

FIGURE 9

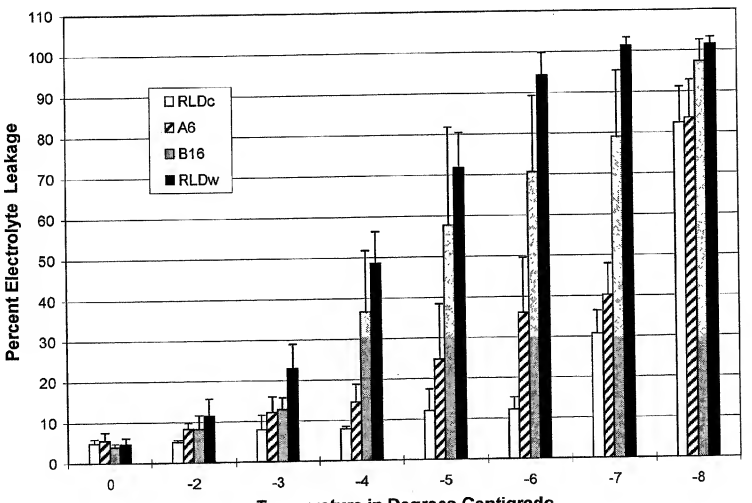

FIGURE 10A

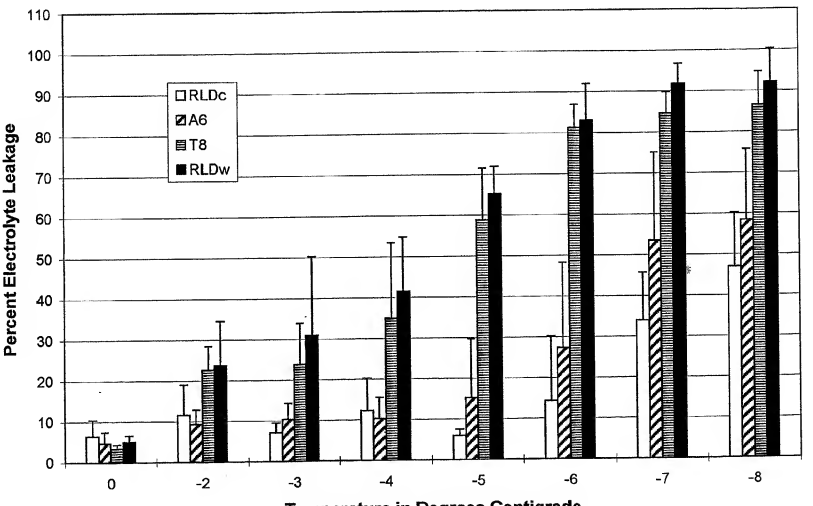

FIGURE 10B

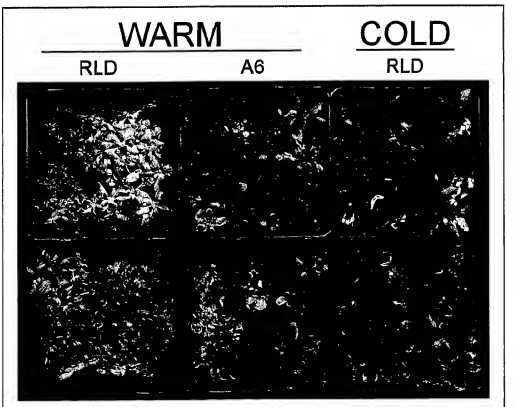

FIGURE 11

```
ATGAACTCATTTTCTGCCTTTTCTGAAATGTTTGGCTCCGATTACGAGTCTCCGGTTTCC  60
Met Asn Ser Phe Ser Ala Phe Ser Glu Met Phe Gly Ser Asp Tyr Glu Ser Pro Val Ser

TCAGGCGGTGATTACAGTCCGAAGCTTGCCACGAGCTGCCCCAAGAAACCAGCGGGAAGG 120
Ser Gly Gly Asp Tyr Ser Pro Lys Leu Ala Thr Ser Cys Pro Lys Lys Pro Ala Gly Arg

AAGAAGTTTCGTGAGACTCGTCACCCAATTTACAGAGGAGTTCGTCAAAGAAACTCCGGT 180
Lys Lys Phe Arg Glu Thr Arg His Pro Ile Tyr Arg Gly Val Arg Gln Arg Asn Ser Gly

AAGTGGGTGTGTGAGTTGAGAGAGCCAAACAAGAAAACGAGGATTTGGCTCGGGACTTTC 240
Lys Trp Val Cys Glu Leu Arg Glu Pro Asn Lys Lys Thr Arg Ile Trp Leu Gly Thr Phe

CAAACCGCTGAGATGGCAGCTCGTGCTCACGACGTCGCCGCCATAGCTCTCCGTGGCAGA 300
Gln Thr Ala Glu Met Ala Ala Arg Ala His Asp Val Ala Ala Ile Ala Leu Arg Gly Arg

TCTGCCTGTCTCAATTTCGCTGACTCGGCTTGGCGGCTACGAATCCCGGAATCAACCTGT 360
Ser Ala Cys Leu Asn Phe Ala Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Ser Thr Cys

GCCAAGGAAATCCAAAAGGCGGCGGCTGAAGCCGCGTTGAATTTTCAAGATGAGATGTGT 420
Ala Lys Glu Ile Gln Lys Ala Ala Ala Glu Ala Ala Leu Asn Phe Gln Asp Glu Met Cys

CATATGACGACGGATGCTCATGGTCTTGACATGGAGGAGACCTTGGTGGAGGCTATTTAT 480
His Met Thr Thr Asp Ala His Gly Leu Asp Met Glu Glu Thr Leu Val Glu Ala Ile Tyr

ACGCCGGAACAGAGCCAAGATGCGTTTTATATGGATGAAGAGGCGATGTTGGGGATGTCT 540
Thr Pro Glu Gln Ser Gln Asp Ala Phe Tyr Met Asp Glu Glu Ala Met Leu Gly Met Ser

AGTTTGTTGGATAACATGGCCGAAGGGATGCTTTTACCGTCGCCGTCGGTTCAATGGAAC 600
Ser Leu Leu Asp Asn Met Ala Glu Gly Met Leu Leu Pro Ser Pro Ser Val Gln Trp Asn

TATAATTTTGATGTCGAGGGAGATGATGACGTGTCCTTATGGAGCTATTAA 651
Tyr Asn Phe Asp Val Glu Gly Asp Asp Asp Val Ser Leu Trp Ser Tyr  •
```

FIGURE 12

```
ATGAACTCATTTTCTGCTTTTTCTGAAATGTTTGGCTCCGATTACGAGTCTTCGGTTTCC 60
Met Asn Ser Phe Ser Ala Phe Ser Glu Met Phe Gly Ser Asp Tyr Glu Ser Ser Val Ser

TCAGGCGGTGATTATATTCCGACGCTTGCGAGCAGCTGCCCCAAGAAACCGGCGGGTCGT 120
Ser Gly Gly Asp Tyr Ile Pro Thr Leu Ala Ser Ser Cys Pro Lys Lys Pro Ala Gly Arg

AAGAAGTTTCGTGAGACTCGTCACCCAATATACAGAGGAGTTCGTCGGAGAAACTCCGGT 180
Lys Lys Phe Arg Glu Thr Arg His Pro Ile Tyr Arg Gly Val Arg Arg Arg Asn Ser Gly

AAGTGGGTTTGTGAGGTTAGAGAACCAAACAAGAAAACAAGGATTTGGCTCGGAACATTT 240
Lys Trp Val Cys Glu Val Arg Glu Pro Asn Lys Lys Thr Arg Ile Trp Leu Gly Thr Phe

CAAACCGCTGAGATGGCAGCTCGAGCTCACGACGTTGCCGCTTTAGCCCTTCGTGGCCGA 300
Gln Thr Ala Glu Met Ala Ala Arg Ala His Asp Val Ala Ala Leu Ala Leu Arg Gly Arg

TCAGCCTGTCTCAATTTCGCTGACTCGGCTTGGAGACTCCGAATCCCGGAATCAACTTGC 360
Ser Ala Cys Leu Asn Phe Ala Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Ser Thr Cys

GCTAAGGACATCCAAAAGGCGGCGGCTGAAGCTGCGTTGGCGTTTCAGGATGAGATGTGT 420
Ala Lys Asp Ile Gln Lys Ala Ala Ala Glu Ala Ala Leu Ala Phe Gln Asp Glu Met Cys

GATGCGACGACGGATCATGGCTTCGACATGGAGGAGACGTTGGTGGAGGCTATTTACACG 480
Asp Ala Thr Thr Asp His Gly Phe Asp Met Glu Glu Thr Leu Val Glu Ala Ile Tyr Thr

GCGGAACAGAGCGAAAATGCGTTTTATATGCACGATGAGGCGATGTTTGAGATGCCGAGT 540
Ala Glu Gln Ser Glu Asn Ala Phe Tyr Met His Asp Glu Ala Met Phe Glu Met Pro Ser

TTGTTGGCTAATATGGCAGAAGGGATGCTTTTGCCGCTTCCGTCCGTACAGTGGAATCAT 600
Leu Leu Ala Asn Met Ala Glu Gly Met Leu Leu Pro Leu Pro Ser Val Gln Trp Asn His

AATCATGAAGTCGACGGCGATGATGACGACGTATCGTTATGGAGTTATTAA 651
Asn His Glu Val Asp Gly Asp Asp Asp Asp Val Ser Leu Trp Ser Tyr •
```

FIGURE 13

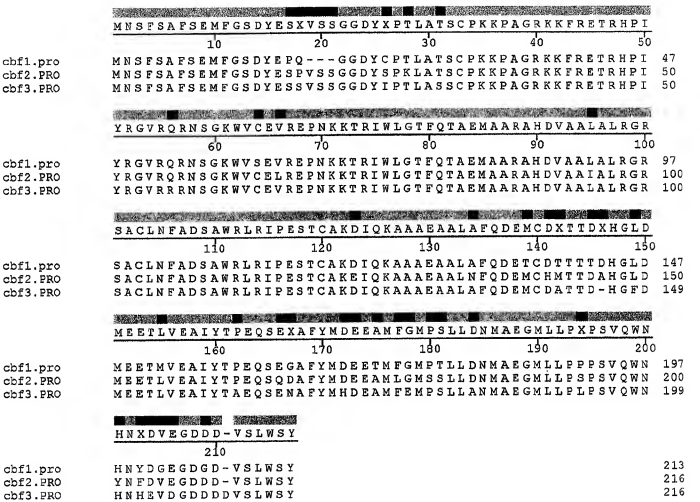

FIGURE 14

CBF1, CBF2 and CBF3 Activity in CRT/DRE Yeast Reporter Strains

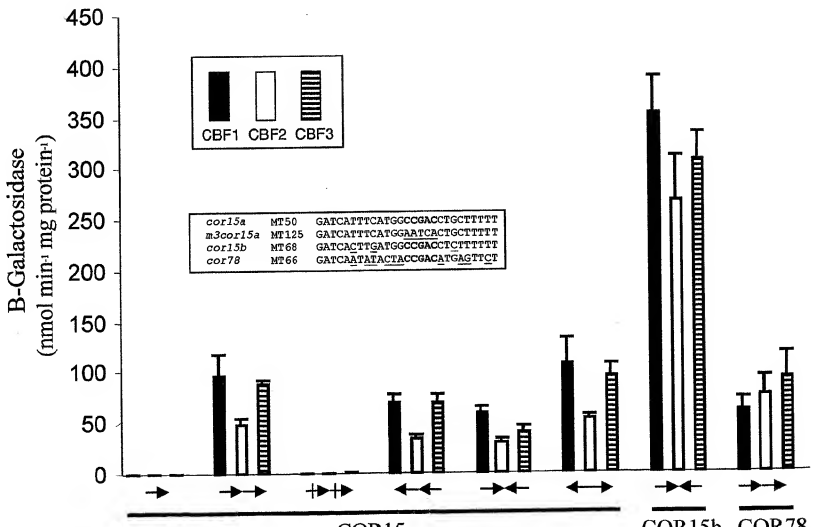

FIGURE 15

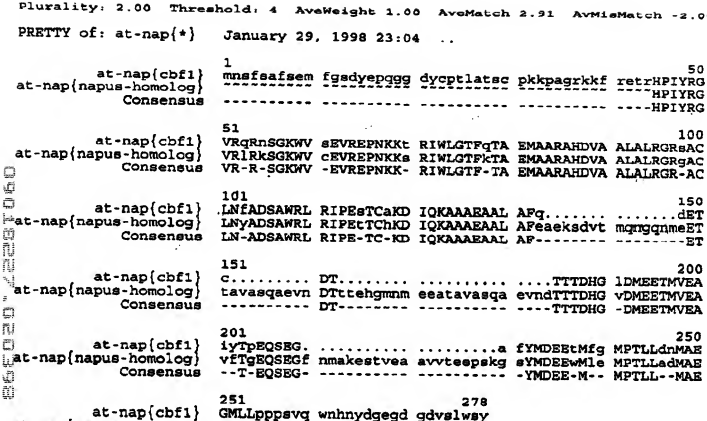
Plurality: 2.00 Threshold: 4 AveWeight 1.00 AveMatch 2.91 AvMisMatch -2.0
PRETTY of: at-nap{*} January 29, 1998 23:04 ..
1 50
at-nap{cbf1} mnsfsafsem fgsdyepggg dycptlatsc pkkpagrkkf retrHPIYRG
at-nap{napus-homolog} ---------- ---------- ---------- ---------- ----HPIYRG
Consensus ---------- ---------- ---------- ---------- ----HPIYRG
51 100
at-nap{cbf1} VRqRnSGKWV sEVREPNKKt RIWLGTFqTA EMAARAHDVA ALALRGRsAC
at-nap{napus-homolog} VRlRkSGKWV cEVREPNKKs RIWLGTFkTA EMAARAHDVA ALALRGRgAC
Consensus VR-R-SGKWV -EVREPNKK- RIWLGTF-TA EMAARAHDVA ALALRGR-AC
101 150
at-nap{cbf1} LNfADSAWRL RIPEsTCaKD IQKAAAEAAL AFq..... .....dET
at-nap{napus-homolog} LNyADSAWRL RIPEtTChKD IQKAAAEAAL AFeaeksdvt mqmggqmmeET
Consensus LN-ADSAWRL RIPE-TC-KD IQKAAAEAAL AF-------- --------ET
151 200
at-nap{cbf1} c..... DT..... ....TTTDHG lDMEETMVEA
at-nap{napus-homolog} tavasqaevn DTttehgmnm eeatavasqa evndTTTDHG vDMEETMVEA
Consensus ---------- DT-------- ---------- ----TTTDHG -DMEETMVEA
201 250
at-nap{cbf1} iyTpEQSEG. ....a fYMDEEtMfg MPTLLdnMAE
at-nap{napus-homolog} vfTgEQSEGf nmakestvea avvteepskg sYMDEEwMle MPTLLadMAE
Consensus --T-EQSEG- ---------- ---------- -YMDEE-M-- MPTLL--MAE
251 278
at-nap{cbf1} GMLLpppsvg wnhnydgegd gdvslwsy
at-nap{napus-homolog} GMLL------ ---------- --------
Consensus GMLL------ ---------- --------

FIGURE 16